

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2001, 14:29:17 ; Search time 15.97 Seconds
(without alignments)
1645.599 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 345
Sequence: 1 MSLEGLLLLTALAGROGT.....DVALEHHEECDCVCRGTGG 345

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 219241 seqs, 76174552 residues

Word size : 0
Total number of hits satisfying chosen parameters: 219241
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	2.3	221	2 D82294	DNA mismatch repair
2	8	2.3	330	2 S60211	fomB protein - Str
3	8	2.3	380	2 A83458	hypothetical prote
4	8	2.3	461	2 A82220	hypothetical prote
5	8	2.3	496	2 F75257	hypothetical prote
6	8	2.3	707	2 JC2218	procollagen C-endo
7	8	2.3	730	1 BMH01	procollagen C-endo
8	8	2.3	823	1 A58788	procollagen C-endo
9	8	2.3	886	1 B58788	procollagen C-endo
10	8	2.3	991	2 I49540	procollagen C-endo
11	7	2.0	57	2 I47032	integrin alpha 5 -
12	7	2.0	68	2 I49136	dopamine transport
13	7	2.0	98	2 B29846	probable activator
14	7	2.0	135	2 G72714	hypothetical prote
15	7	2.0	136	2 S69638	ribosomal protein
16	7	2.0	138	2 S51556	K222 protein - hum
17	7	2.0	148	2 S46044	hypothetical prote
18	7	2.0	177	2 E71211	hypothetical prote
19	7	2.0	184	2 C70242	conserved hypothet
20	7	2.0	185	2 S74428	probable pseudouri
21	7	2.0	222	2 S61241	hypothetical prote
22	7	2.0	239	2 B81345	hypothetical prote
23	7	2.0	245	2 S28827	probable periplasm
24	7	2.0	250	2 T08680	chlorophyll a/b-bi
25	7	2.0	254	2 T29138	hypothetical prote
26	7	2.0	258	2 D81127	hypothetical prote
27	7	2.0	283	2 F82108	probable ferredoxi
28	7	2.0	283	2 C83248	2-dehydro-3-deoxyp
29	7	2.0	296	2 S51746	hypothetical prote

30 7 2.0 299 2 T34987 probable integral
31 7 2.0 299 2 T08704 hypothetical prote
32 7 2.0 313 2 E69822 hypothetical prote
33 7 2.0 319 2 B83560 probable transmem
34 7 2.0 320 2 H85489 hypothetical prote
35 7 2.0 321 2 A82070 transcription regu
36 7 2.0 338 2 JC5707 HVA22 protein - hu
37 7 2.0 340 2 JC7505 brain link protein
38 7 2.0 340 2 C70030 hypothetical prote
39 7 2.0 345 2 T17053 NADH dehydrogenase
40 7 2.0 345 2 T17062 NADH dehydrogenase
41 7 2.0 345 2 JC4665 protein kinase (EC
42 7 2.0 352 2 T06463 phosphoribulokinas
43 7 2.0 368 2 E83003 hypothetical prote
44 7 2.0 373 2 D64729 probable activator
45 7 2.0 374 2 A83761 hypothetical prote

ALIGNMENTS

RESULT 1

DNA mismatch repair protein Muth VC0668 [imported] - Vibrio cholerae (strain N16961 s
D82294
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82294
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833

A:Accession: D82294

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-221 <HEI>

A:Cross-references: GB:AF004153; GB:AE003852; NID:g9655103; PIDN:AAF93833.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0668

A:Map position: 1

C:Superfamily: mutator muth

Query Match 2.3%; Score 8; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 307 EVLQLRPK 314
|||||||

Db 173 EVLQLRPK 180

RESULT 2

S60211

fomB protein - Streptomyces wedmorensis

C:Species: Streptomyces wedmorensis

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C:Accession: S60211

R:Hidaka, T.; Goda, M.; Kuzuyama, T.; Takei, N.; Hidaka, M.; Seto, H.

Mol. Gen. Genet. 249, 274-280, 1995

A:Title: Cloning and nucleotide sequence of fosfomycin biosynthetic genes of Streptom

A:Reference number: S60207; MUID:96091152

A:Accession: S60211

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-330 <HID>

A:Cross-references: EMBL:D39561; NID:g3452578; PIDN:BAA32494.1; PID:g1061006

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: fomB

C:Superfamily: Streptomyces wedmorensis fomB protein

Query Match 2.3%; Score 8; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 SRVVDLNL 240
 |||||
 Db 10 SRVVDLNL 17

RESULT 3
 A:3458
 hypothetical protein PA1509 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: A83458
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337
 A:Accession: A83458
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-380 <STO>
 A:Cross-references: GB:AE004579; GB:AE004091; NID:g9947455; PIDN:AAG04898.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1509

Query Match 2.3%; Score 8; DB 2; Length 380;
 Best Local Similarity 100.0%; Pred. No. 6.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 PSEPGFCI 158
 |||||
 Db 174 PSEPGFCI 181

RESULT 4
 A:82220
 hypothetical protein VC1265 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: A82220
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833
 A:Accession: A82220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <HEI>
 A:Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94424.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1265
 A:Map position: 1

Query Match 2.3%; Score 8; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLLLSAL 13
 |||||
 Db 7 LLLLSAL 14

RESULT 5
 F75257
 hypothetical protein - Deinococcus radiodurans (strain RL)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: F75257
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
 .; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL.
 A:Reference number: A75250; MUID:20036896
 A:Accession: F75257
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-496 <WHI>
 A:Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF12116.1; PID:g646
 A:Experimental source: strain RL
 C:Genetics:
 A:Gene: DR2572
 A:Map position: 1

Query Match 2.3%; Score 8; DB 2; Length 496;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLLLSAL 13
 |||||
 Db 304 LLLLSAL 311

RESULT 6
 JC2218
 procollagen C-endopeptidase (EC 3.4.24.19) - African clawed frog (clone 22AN)
 N:Alternate names: bone morphogenic protein 1
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Sep-1999
 C:Accession: JC2218
 R:Maeno, M.; Xue, Y.; Wood, T.I.; Ong, R.C.; Kung, H.
 Gene 134, 257-261, 1993
 A:Title: Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic pr
 A:Reference number: JC2218; MUID:94085787
 A:Accession: JC2218
 A:Molecule type: mRNA
 A:Residues: 1-707 <NAE>
 A:Cross-references: GB:L12249; NID:g406540; PIDN:AAA16313.1; PID:g406541
 C:Comment: This protein induces ectopic cartilage formation in vivo.
 C:Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology
 C:Keywords: beta-hydroxyasparagine; glycoprotein; hydrolase; metalloproteinase; zinc
 F:93-284/Domain: astacin homology <AST>
 F:285-397/Region: complement lr/ls-like repeat
 F:285-394/Domain: Clr/Cls repeat homology <ClrI>
 F:398-510/Region: complement lr/ls-like repeat
 F:398-507/Domain: Clr/Cls repeat homology <Clr2>
 F:514-550/Domain: EGF homology <EGF>
 F:554-666/Region: complement lr/ls-like repeat
 F:554-663/Domain: Clr/Cls repeat homology <ClrI>
 F:62,105,295,326/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:176,180,186,235/Binding site: zinc (His, His, Tyr) #status predicted
 F:177/Active site: Glu #status predicted
 F:528/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 2.3%; Score 8; DB 2; Length 707;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 CKYDFVEV 111
 |||||
 Db 607 CKYDFVEV 614

RESULT 7

BMH1
procollagen C-endopeptidase (EC 3.4.24.19) precursor, splice form BMP1 - human
N:Alternate names: bone morphogenic protein 1 (BMP1)
C:Species: Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
C:Accession: A37278; E58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278
A:Molecule type: mRNA
A:Residues: 1-730 <WOZ>
A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500
C:Genetics:
A:Gene: GDB:BMPI
A:Cross-references: GDB:125203; OMIM:112264
A:Map position: 8p21-8p21

C:Function:
A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology;
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; g
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: procollagen C-endopeptidase splice form BMP1 #status predicted <MAT>
F:130-321/Domain: astacin homology <AST>
F:435-544/Domain: C1r/C1s repeat homology <C1R1>
F:551-587/Domain: EGF homology <EGF>
F:591-700/Domain: C1r/C1s repeat homology <C1R2>
F:738-752/Region: histidine-rich
F:91,142,332,363,599/Binding site: carboxydrate (Asn) (covalent) #status predicted
F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,559-572,574-587,591-617,644
F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 2.3%; Score 8; DB 1; Length 730;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111

|||||

Db 644 CKYDFVEV 651

RESULT 8

procollagen C-endopeptidase (EC 3.4.24.19) precursor, splice form HIS - human
N:Alternate names: bone morphogenic protein splice form BMP-1/HIS
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 18-Jun-1999
C:Accession: A37278; A58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278

A:Molecule type: mRNA

A:Residues: 1-702, 'EKRPALQPPGRPHQLKFRVQKRNTPQ' <WOZ>

A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500

R:Takahara, K.; Lyons, G.E.; Greenspan, D.S.

J. Biol. Chem. 269, 32572-32578, 1994

A:Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encod

A:Reference number: A58788; MUID:95096114

A:Accession: A58788

A:Molecule type: mRNA

A:Residues: 703-823 <TAK>

A:Cross-references: GB:L35278; NID:g619423; PIDN:AAC41703.1; PID:g619424

C:Genetics:

A:Gene: GDB:BMPI; BMP-1

A:Cross-references: GDB:125203; OMIM:112264

A:Map position: 8p21-8p21

C:Function:

A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen t
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology;
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication;
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-823/Product: procollagen C-endopeptidase splice form HIS #status predicted <MAT>

F:130-321/Domain: astacin homology <AST>

F:322-431/Domain: C1r/C1s repeat homology <C1R1>

F:435-544/Domain: C1r/C1s repeat homology <C1R2>

F:551-587/Domain: EGF homology <EGF>

F:591-700/Domain: C1r/C1s repeat homology <C1R3>

F:738-752/Region: histidine-rich

F:91,142,332,363,599/Binding site: carboxydrate (Asn) (covalent) #status predicted

F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,559-572,574-587,591-617,644

F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted

F:214/Active site: Glu #status predicted

F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 2.3%; Score 8; DB 1; Length 823;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111

|||||

Db 644 CKYDFVEV 651

RESULT 9

procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - huma
N:Alternate names: bone morphogenic protein 1, tolloid-like splice form
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 16-Jul-1999
C:Accession: A37278; B58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.;
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278

A:Molecule type: mRNA

A:Residues: 1-702, 'EKRPALQPPGRPHQLKFRVQKRNTPQ' <WOZ>

A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500

R:Takahara, K.; Lyons, G.E.; Greenspan, D.S.

J. Biol. Chem. 269, 32572-32578, 1994

A:Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are en

A:Reference number: A58788; MUID:95096114

A:Accession: B58788

A:Molecule type: mRNA

A:Residues: 703-986 <TAK>

A:Cross-references: GB:L35279; NID:g619860; PIDN:AAC41710.1; PID:g619861

C:Genetics:

A:Gene: GDB:BMPI; BMP-1

A:Cross-references: GDB:125203; OMIM:112264

A:Map position: 8p21-8p21

C:Function:

A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen t
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology;
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication;
F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-986/Product: procollagen C-endopeptidase tolloid-like splice form #status predic

F:130-321/Domain: astacin homology <AST>

F:322-431/Domain: C1r/C1s repeat homology <C1R1>

F:435-544/Domain: C1r/C1s repeat homology <C1R2>

F:551-587/Domain: EGF homology <EGF>

F:591-700/Domain: C1r/C1s repeat homology <C1R3>

F:707-742/Domain: EGF homology <EG2>

F:747-856/Domain: C1r/C1s repeat homology <C1R4>

F:860-973/Domain: C1r/C1s repeat homology <C1R5>

F:91,142,332,363,599/Binding site: carboxydrate (Asn) (covalent) #status predicted

F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,559-572,574-587,591-617,644

F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted

F:214/Active site: Glu #status predicted
F:565,720/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 2.3%; Score 8; DB 1; Length 986;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111
|||||||
Db 644 CKYDFVEV 651

RESULT 10
I49540
procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: I49540
R:Fukagawa, M.; Noboru, S.; Hagan, B.L.M.; Jones, C.M.
Dev. Biol. 163, 175-183, 1994
A:Title: Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1) which is related to the Drosophila gene *decapentagene*
A:Reference number: I49540; MUID:94229342
A:Accession: I49540
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-991 <RES>
A:Cross-references: GB:L24755; NID:g439606; PIDN:AAA37306.1; PID:g439607
C:Genetics:
A:Gene: Bmp-1
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology; H

Query Match 2.3%; Score 8; DB 2; Length 991;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111
|||||||
Db 649 CKYDFVEV 656

RESULT 11
I47032
integrin alpha 5 - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 29-Sep-1999
C:Accession: I47032
R:Ohashi, H.; Maeda, T.; Mishima, H.; Otori, T.; Nishida, T.; Sekiguchi, K.
Exp. Cell Res. 219, 418-423, 1995
A:Title: Up-regulation of integrin alpha 5 beta 1 expression by interleukin-6 in rabbit
A:Reference number: I47032; MUID:95317375
A:Accession: I47032
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-57 <ORF>
A:Cross-references: GB:S77513; NID:g957336; PIDN:AAB34683.1; PID:g957337
C:Superfamily: integrin alpha-2b chain

Query Match 2.0%; Score 7; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LFGLLLL 9
|||||||

Db 15 LFGLLLL 21

RESULT 12
I49136
dopamine transporter - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: I49136
R:Donovan, D.M.; Vandenbergh, D.J.; Perry, M.P.; Bird, G.S.; Ingersoll, R.; Nanthakum
Brain Res. Mol. Brain Res. 30, 327-335, 1995
A:Title: Human and mouse dopamine transporter genes: conservation of 5'-flanking sequence
A:Reference number: I37296; MUID:95364623
A:Accession: I49136
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-68 <RES>
A:Cross-references: EMBL:U16265; NID:g1055215; PIDN:AAC52283.1; PID:g1055216
C:Genetics:
A:Gene: DAT
C:Superfamily: gamma-aminobutyric acid transporter

Query Match 2.0%; Score 7; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KEQGVQ 42
|||||||
Db 35 KEQGVQ 41

RESULT 13
B29846
probable activator protein leuO - Salmonella typhimurium (fragment)
C:Species: Salmonella typhimurium
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Sep-1999
C:Accession: B29846; S78513
R:Haughn, G.W.; Wessler, S.R.; Gemmill, R.M.; Calvo, J.M.
J. Bacteriol. 166, 1113-1117, 1986
A:Title: High A+T content conserved in DNA sequences upstream of leuABCD in Escherichia coli
A:Reference number: A91815; MUID:86223773
A:Accession: B29846
A:Molecule type: DNA
A:Residues: 1-98 <HAU>
R:Haughn, G.W.; Wessler, S.R.; Gemmill, R.M.; Calvo, J.M.
submitted to the EMBL Data Library, June 1996
A:Reference number: S78513
A:Accession: S78513
A:Molecule type: DNA
A:Residues: 37-98 <HAG>
A:Cross-references: EMBL:M12892; NID:g154162; PIDN:AAB02429.1; PID:g1374958
C:Genetics:
A:Gene: leuO
A:Map position: 43 min
C:Superfamily: leuO protein
C:Keywords: DNA binding; transcription regulation
F:39-58/Region: helix-turn-helix motif
F:61-91/Region: regulatory protein lysR motif

Query Match 2.0%; Score 7; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 VDLNLLT 242
|||||||
Db 45 VDLNLLT 51

RESULT 14
G72714

Search completed: August 22, 2001, 14:32:01
Job time: 164 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2001, 14:30:17 ; Search time 24.64 Seconds
(without alignments)
1852.485 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 345
Sequence: 1 MSFLGLLLLSALAGROQT.....DVALEHHEECDCVCRGSGG 345

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL16.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_vertebrate.*
14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	345	100.0	345	4 Q9UL22	Q9ul22 homo sapien
2	327	94.8	345	4 Q9NRA1	Q9nra1 homo sapien
3	52	15.1	345	11 Q9QY71	Q9qy71 mus musculus
4	52	15.1	345	11 Q9EOX6	Q9eqx6 rattus norv
5	37	10.7	345	13 Q9I946	Q9i946 gallus gall
6	31	9.0	345	11 Q9JHV8	Q9jvh8 mus musculus
7	8	2.3	116	10 P93284	P93284 arabidopsis
8	8	2.3	221	2 Q9KU56	Q9ku56 vibrio chol
9	8	2.3	241	11 Q9Z135	Q9z135 rattus norv
10	8	2.3	292	4 Q9H9G2	Q9h9g2 homo sapien
11	8	2.3	330	2 Q56188	Q56188 streptomyce
12	8	2.3	370	4 Q9GZP0	Q9gzp0 homo sapien
13	8	2.3	380	2 Q9I3K3	Q9i3k3 pseudomonas
14	8	2.3	461	2 Q9KSJ1	Q9ksj1 vibrio chol
15	8	2.3	496	2 Q9RRC0	Q9rrc0 deinococcus
16	8	2.3	558	2 Q9WXB0	Q9wxb0 acidiphiliu
17	8	2.3	691	13 Q57658	Q57658 gallus gall
18	8	2.3	735	13 Q57381	Q57381 xenopus lae
19	8	2.3	886	2 Q05560	Q05560 mycobacteri

Query Match 100.0%; Score 345; DB 4; Length 345;

20	8	2.3	926	4 Q9UQ00	Q9uq00 homo sapien
21	8	2.3	977	13 Q91925	Q91925 xenopus lae
22	8	2.3	985	4 Q75184	Q75184 homo sapien
23	8	2.3	1012	11 Q9WVW6	Q9wvwm mus musculu
24	8	2.3	1015	4 Q9Y6L7	Q9y6l7 homo sapien
25	8	2.3	1081	4 Q43485	Q43485 homo sapien
26	8	2.3	1099	4 Q60721	Q60721 homo sapien
27	8	2.3	1837	5 Q9VCZ9	Q9vcz9 drosophila
28	7	2.0	57	6 Q28897	Q28897 oryctolagus
29	7	2.0	95	4 Q9Y610	Q9y610 homo sapien
30	7	2.0	135	1 Q9YCX8	Q9ycx8 aeropyrum p
31	7	2.0	138	4 Q14867	Q14867 homo sapien
32	7	2.0	138	4 Q9H7R1	Q9h7r1 homo sapien
33	7	2.0	142	11 Q9RIA1	Q9ria1 mus musculu
34	7	2.0	148	8 Q9G9S8	Q9gs98 metridia lu
35	7	2.0	150	2 Q9X5L7	Q9x5l7 rhodobacter
36	7	2.0	153	4 Q9NRE7	Q9nre7 homo sapien
37	7	2.0	177	1 Q59617	Q59617 pyrococcus
38	7	2.0	184	2 Q50874	Q50874 borrelia bu
39	7	2.0	190	5 Q9NK67	Q9nk67 drosophila
40	7	2.0	207	10 P94014	P94014 arabidopsis
41	7	2.0	208	10 P94097	P94097 arabidopsis
42	7	2.0	222	14 Q65570	Q65570 bovine herp
43	7	2.0	223	5 Q9VJM2	Q9vjm2 drosophila
44	7	2.0	230	14 Q9QAB2	Q9qab2 neodiprion
45	7	2.0	239	2 Q9PPH2	Q9pph2 campylobact

ALIGNMENTS

RESULT 1

Q9UL22 ID Q9UL22 PRELIMINARY; PRT; 345 AA.
AC Q9UL22;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SECRETORY GROWTH FACTOR-LIKE PROTEIN FALLOTEIN (SPINAL CORD-DERIVED GROWTH FACTOR).
GN HSCDGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Tsai Y.J., Lee R.K.K., Lin S.P.;
RT "Fallotein, a novel growth factor like gene identified in human uterus."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
MEDLINE=20317014; PubMed=10858496;
RA Hamada T., Ui-Tei K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDGF, is a unique member of the PDGF/VEGF family."
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AF091434; AAF00049.1;
DR EMBL; AB033831; BAB03266.1;
DR InterPro; IPR000072;
DR InterPro; IPR008859;
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39029 MW; CDE9E51F40633E78 CRC64;

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTALAGORCTOAESNLSSKFQSSNKQNGVODPOHERIITVSTNGSIHS 60
Db 1 MSFLGILLTALAGORCTOAESNLSSKFQSSNKQNGVODPOHERIITVSTNGSIHS 60
QY 61 PRPHYTPRNTVLVRLVAENWVLIQTFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120
Db 61 PRPHYTPRNTVLVRLVAENWVLIQTFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSGTVPGKQISKGNQIRIRFVSEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSSA 180
Db 121 GRWCGSGTVPGKQISKGNQIRIRFVSEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSSA 180
QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNL 240
Db 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNL 240
QY 241 LTEEVRLYSCTPRNFSVSIRELKRDTIFWPGCLLVKRCGNCACCLHNCNECQCVPK 300
Db 241 LTEEVRLYSCTPRNFSVSIRELKRDTIFWPGCLLVKRCGNCACCLHNCNECQCVPK 300
QY 301 VTKYHEVQLRPTKGTGRLHKSLLTDVALEHHECDCVCRGSTG 345
Db 301 VTKYHEVQLRPTKGTGRLHKSLLTDVALEHHECDCVCRGSTG 345

RESULT 2
Q9NRAL
ID Q9NRAL PRELIMINARY; PRT; 345 AA.
AC Q9NRAL;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Li X., Ponten A., Aase K., Karlsson L., Abramsson A., Uutela M.,
RA Backstrom G., Hellstrom M., Bostrom H., Li H., Soriano P.,
RA Betsholtz C., Heldin C.-H., Alltalo K., Osman A., Eriksson U.;
RT "PDGF-C is a novel protease-activated ligand for the PDGF alpha
RT receptor."
RL Nat. Cell Biol. 0:0-0(2000).
DR EMBL: AF244813; AAF80597.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39043 MW; 590889CEA55CC5EA CRC64;

Query Match 94.8%; Score 327; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GTQAESNLSSKFQSSNKQNGVODPOHERIITVSTNGSIHSIPRPHYTPRNTVLVRLV 78
Db 19 GTQAESNLSSKFQSSNKQNGVODPOHERIITVSTNGSIHSIPRPHYTPRNTVLVRLV 78
QY 79 AVEENWVLIQTFDERFGLDEPDDICKYDFVEVEEPSDGTILGRWCGSGTVPGKQISKGN 138
Db 79 AVEENWVLIQTFDERFGLDEPDDICKYDFVEVEEPSDGTILGRWCGSGTVPGKQISKGN 138
QY 159 QIRIRFVSEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSSALPDLNNAITAFSTLED 198
```

```
|||||
Db 139 QIRIRFVSEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSSALPDLNNAITAFSTLED 198
QY 199 LIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNLTEEVRLYSCTPRNFSVS 258
Db 199 LIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNLTEEVRLYSCTPRNFSVS 258
QY 259 IREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPKSKYTKYHEVQLRPTKGTGR 318
Db 259 IREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPKSKYTKYHEVQLRPTKGTGR 318
QY 319 GLHKSLLTDVALEHHECDCVCRGSTG 345
Db 319 GLHKSLLTDVALEHHECDCVCRGSTG 345

RESULT 3
Q9QY71
ID Q9QY71 PRELIMINARY; PRT; 345 AA.
AC Q9QY71;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FALLOTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Tsai Y.-J., Lee R.K.-K., Chen Y.-H., Lin S.-P., Cheng W.T.-K.;
RT "cDNA cloning of follotein from mouse ovary."
RL Submitted (Jan-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117608; AAF22516.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38741 MW; 3A58A1F701B84EA2 CRC64;

Query Match 15.1%; Score 52; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 3.5e-44;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 LXSCTPRNFSVSIRELKRDTIFWPGCLLVKRCGNCACCLHNCNECQCVP 298
Db 247 LXSCTPRNFSVSIRELKRDTIFWPGCLLVKRCGNCACCLHNCNECQCVP 298

RESULT 4
Q9EQX6
ID Q9EQX6 PRELIMINARY; PRT; 345 AA.
AC Q9EQX6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN RSCDGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=KIDNEY;
RA Hamada T., Ui-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDGF-B, a Novel Growth Factor Homologous to
RT SCDGF/PDGF-C/fallotein."
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
```

DR EMBL; AB033830; BAB19969.1; --
SQ SEQUENCE 345 AA; 38734 MW; F296DA6B9B765D10 CRC64;

Query Match 15.1%; Score 52; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 3.5e-44;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 LYSCTPRNFSVSIRELKRTDTTFWPGCLLVKRCGGNCACCLHNCNCQVCP 298
DB 247 LYSCTPRNFSVSIRELKRTDTTFWPGCLLVKRCGGNCACCLHNCNCQVCP 298

RESULT 5
Q91946 PRELIMINARY; PRT; 345 AA.
ID Q91946;
AC Q91946;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN SCDFG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=SPINAL CORD;
RX MEDLINE=20317014; PubMed=10858496;
RA Hamada T., Ui-Tel K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDFG, is a unique member of the PDGF/VEGF family."
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AB033829; BAB03265.1; --
DR InterPro; IPR000072; --
DR InterPro; IPR000859; --
Qr Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38940 MW; 97ACEA992BF5128C CRC64;

Query Match 10.7%; Score 37; DB 13; Length 345;
Best Local Similarity 100.0%; Pred. No. 5.6e-29;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 ENWVIQTFDEREGLEDDICKYDFVEVEEPSDGT 118
DB 82 ENWVIQTFDEREGLEDDICKYDFVEVEEPSDGT 118

RESULT 6
Q9JHV8 PRELIMINARY; PRT; 345 AA.
ID Q9JHV8;
AC Q9JHV8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
GN PDGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS-WEBSTER/NIH;
RA Ding H., Wu X., Kim I., Tam P.P.L., Koh G.Y., Nagy A.;
RT "The mouse pdgfc gene: Dynamic expression in embryonic tissues during organogenesis.";

RL Mech. Dev. 0:0-0(2000).
DR EMBL; AF286725; AAF91483.1; --
DR InterPro; IPR000072; --
DR InterPro; IPR000859; --
Qr Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38886 MW; FA1486BED6D362F8 CRC64;

Query Match 9.0%; Score 31; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 6.7e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 QLTDFEREGLEDDICKYDFVEVEEPSDG 117
DB 87 QLTDFEREGLEDDICKYDFVEVEEPSDG 117

RESULT 7
P93284 PRELIMINARY; PRT; 116 AA.
ID P93284;
AC P93284;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-NOV-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE ORF116.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Unseld M., Marienfeld J.R., Brandt P., Brennicke A.;
RL Nat. Genet. 0:0-0(0).
DR EMBL; Y08501; CAA69760.1; --
DR Mendel; 13039; Arath; 2239; 13039.
SQ SEQUENCE 116 AA; 13352 MW; 98BEEFAF9F5FC2BC CRC64;

Query Match 2.3%; Score 8; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLLTSL 13
DB 63 LLLTSL 70

RESULT 8
Q9KU56 PRELIMINARY; PRT; 221 AA.
ID Q9KU56;
AC Q9KU56;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MUTH.
GN VC0668.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RL *cholerae*.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004153; AAF93833.1; -.
 DR TIGR: VC0668; -.
 SQ SEQUENCE 221 AA; 24584 MW; 12A8F351F3E5D051 CRC64;

Query Match 2.3%; Score 8; DB 2; Length 221;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 307 EVLQLRPK 314
 |||||
 Db 173 EVLQLRPK 180

RESULT 9
 ID Q92135 PRELIMINARY; PRT; 241 AA.
 AC Q92135;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAR-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE PROCOLLAGEN C-PROTEINASE 3 (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 SEQUENCE FROM N.A.
 RN Rattus-TESTIS;
 RA Hirahara I., Tomita M., Uneyama K., Urakami K.;
 RT "New splicing pattern of the gene for procollagen C-proteinase.";
 RL Cell Struct. Funct. 23:125-125(1998).
 DR EMBL: AB012139; BAA75639.1; -.
 DR HSP; P00736; IAPQ.
 DR MEROPS: M12.005; -.
 DR InterPro: IPR000152; -.
 DR InterPro: IPR000561; -.
 DR InterPro: IPR000859; -.
 DR InterPro: IPR001881; -.
 DR Pfam: PF000008; EGF; 1.
 DR Pfam: PF00431; CUB; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR SMART: SM00042; CUB; 1.
 KW Calcium-binding; Collagen; EGF-like domain; Glycoprotein;
 KW Hydroxylation; Repeat.
 FT NON_TER 1
 SQ SEQUENCE 241 AA; 27199 MW; 380803EBDE814EFA CRC64;

Query Match 2.3%; Score 8; DB 11; Length 241;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 CKYDFVEV 111
 |||||
 Db 160 CKYDFVEV 167

RESULT 10
 ID Q9H9G2 PRELIMINARY; PRT; 292 AA.
 AC Q9H9G2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CDNA FLJ12781 FTS, CLONE NT2RP2001861.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 RA Isoigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK022843; BAB14267.1; -.
 SQ SEQUENCE 292 AA; 30878 MW; 350F99881A7F6DB1 CRC64;

Query Match 2.3%; Score 8; DB 4; Length 292;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 335 CDCVCRGS 342
 |||||
 Db 19 CDCVCRGS 26

RESULT 11
 Q56188
 ID Q56188 PRELIMINARY; PRT; 330 AA.
 AC Q56188;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE FOMB PROTEIN.
 GN FOMB.
 OS Streptomyces wedmorensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=43759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kuzuyama T.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309717; PubMed=7789803;
 RA Hidaka T., Hidaka M., Kuzuyama T., Seto H.;
 RT "Sequence of a P-methyltransferase-encoding gene isolated from a
 RT bialphos-producing Streptomyces hygroscopicus.";
 RL Gene 158:149-150(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96083774; PubMed=7490235;
 RA Kuzuyama T., Seki T., Dairi T., Hidaka T., Seto H.;
 RT "Nucleotide sequence of fortimicin K11 methyltransferase gene isolated
 RT from Micromonospora olivasterospora, and comparison of its deduced
 RT amino acid sequence with those of methyltransferases involved in the
 RT biosynthesis of bialaphos and fosfomycin.";
 RL J. Antibiot. 48:1191-1193(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96091152; PubMed=7500951;
 RA Hidaka T., Goda M., Kuzuyama T., Takei N., Hidaka M., Seto H.;
 RT "Cloning and nucleotide sequence of fosfomycin biosynthetic genes of
 RT Streptomyces wedmorensis.";
 RL Mol. Gen. Genet. 249:274-280(1995).
 DR EMBL: AB016934; BAA32494.1; -.
 SQ SEQUENCE 330 AA; 36690 MW; C2D1892FD7CC145A CRC64;

Query Match 2.3%; Score 8; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 SRVVDLNL 240
 DB 10 SRVVDLNL 17

RESULT 12
 Q9GZP0 PRELIMINARY; PRT; 370 AA.
 AC Q9GZP0
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE SPINAL CORD-DERIVED GROWTH FACTOR-B (MSTP036).
 GN HSCDGF-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hamada T., Ui-Tei K., Imaki J., Miyata Y.;
 RT "Molecular Cloning of SCDGF-B, a Novel Growth Factor Homologous to
 SCDGF/PDGF-C/fallotelin.";
 RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-AORTA;
 RA Liu B., Liu Y.Q., Wang X.Y., Zhao B., Sheng H., Zhao X.W., Liu S.,
 RA Xu Y.Y., Ye J., Song L., Zhang C.L., Zhang J., Wei Y.J.,
 RA Cao H.Q., Zhao Y., Liu L.S., Ding J.F., Gao R.L., Wu Q.Y., Qiang B.Q.,
 RA Yuan J.G., Liaw C.C., Zhao M.S., Hui R.T.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB033832; BAB18903.1;
 DR EMBL; AF113216; AAG39287.1;
 SQ SEQUENCE 370 AA; 42848 MW; D387F485E7BB7674 CRC64;

Query Match 2.3%; Score 8; DB 4; Length 370;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 KRYHEVIQ 310
 DB 325 KRYHEVIQ 332

RESULT 13
 Q9I3K3 PRELIMINARY; PRT; 380 AA.
 AC Q9I3K3
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL PROTEIN PA1509.
 GN PA1509.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RL Nature 406:959-964(2000).

DR EMBL; AE004579; AAG04898.1;
 KW Hypothetical protein.
 SQ SEQUENCE 380 AA; 42364 MW; 45109F1F66FB9B2A CRC64;

Query Match 2.3%; Score 8; DB 2; Length 380;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PSEPGFCI 158
 DB 174 PSEPGFCI 181

RESULT 14
 Q9KSJ1 PRELIMINARY; PRT; 461 AA.
 AC Q9KSJ1
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL PROTEIN VC1265.
 GN VC1265.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004206; AAF94424.1;
 DR TIGR; VC1265;
 DR InterPro; IPR000345;
 DR PROSITE; PS00190; CVTOCHROME_C; UNKNOWN_2.
 KW Hypothetical protein.
 SQ SEQUENCE 461 AA; 50283 MW; 854ACB61E9D2E4ED CRC64;

Query Match 2.3%; Score 8; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLLLTSL 13
 DB 7 LLLLTSL 14

RESULT 15
 Q9RRCO PRELIMINARY; PRT; 496 AA.
 AC Q9RRCO
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 50.8 KDA PROTEIN.
 GN DR2572.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1;
 RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1";
RL Science 286:1571-1577(1999).
DR EMBL: AE002086; AAF12116.1; -.
DR TIGR: DR2572; -.
KW Hypothetical protein.
SQ SEQUENCE 496 AA; 50789 MW; D3ADCB51F339CDB2 CRC64;

Query Match 2.38; Score 8; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 LLLLTSL 13
| | | | | | | |
Db 304 LLLLTSL 311

Search completed: August 22, 2001, 14:32:32
Job time: 135 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2001, 14:29:11 ; Search time 12.3 Seconds
(without alignments)
577.534 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 345
Sequence: 1.MSLFGLLLLTALACQROGT.....DVALEHHBECDCVCRGSGG 345

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents,AA:*
1: /cgn2.6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2.6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2.6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB ID	Description
		Match	Length		
1	8	2.3	788	1	US-08-572-225-1
2	7	2.0	25	1	US-08-672-571A-5
3	7	2.0	65	1	US-08-426-819A-33
4	7	2.0	248	1	US-08-426-819A-37
5	7	2.0	344	2	US-08-755-728-3
6	7	2.0	344	2	US-08-974-655-3
7	7	2.0	344	4	US-09-283-011-3
8	7	2.0	347	2	US-09-016-000-1
9	7	2.0	456	4	US-09-172-841-51
10	7	2.0	461	1	US-08-672-571A-3
11	7	2.0	490	1	US-08-672-571A-1
12	7	2.0	561	2	US-08-559-492-12
13	7	2.0	584	1	US-08-426-819A-36
14	7	2.0	602	2	US-08-419-652-6
15	7	2.0	616	1	US-07-879-617A-13
16	7	2.0	616	1	US-08-753-985-13
17	7	2.0	619	1	US-07-762-132A-2
18	7	2.0	619	1	US-08-301-722A-4
19	7	2.0	620	1	US-08-301-722A-2
20	7	2.0	622	1	US-08-426-819A-35
21	7	2.0	628	3	US-08-776-271-2
22	7	2.0	628	4	US-09-215-035-2
23	7	2.0	783	6	5422248-2
24	7	2.0	836	1	US-07-923-976-4
25	7	2.0	863	1	US-07-923-976-8
26	6	1.7	10	3	US-08-951-944-18
27	6	1.7	12	3	US-08-997-211-8

US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572/225
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-572-225-1

ALIGNMENTS

RESULT 1
US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572/225
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-572-225-1

Query Match 2.3%; Score 8; DB 1; Length 788;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 CKYDFEV 111
Db 446 CKYDFEV 453

RESULT 2
US-08-672-571A-5
; Sequence 5, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycosylceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
US-08-672-571A-5

Query Match 2.0%; Score 7; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 MPQTEA 170
Db 6 MPQTEA 12

RESULT 3
US-08-426-819A-33
; Sequence 33, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiro
; TITLE OF INVENTION: Genes Coding for Megakaryocyte

; TITLE OF INVENTION: Potentiator
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-426-819A-33

Query Match 2.0%; Score 7; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 ALPLDLL 186
Db 52 ALPLDLL 58

RESULT 4
US-08-426-819A-37
; Sequence 37, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiro
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.

; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
US-08-426-819A-37

Query Match 2.0%; Score 7; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 ALPLDLL 186
|||
Db 79 ALPLDLL 85

RESULT 5

US-08-755-728-3
; Sequence 3, Application US/08755728
; Patent No. 5962312
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,728
; FILING DATE: No. 5962312ember 25, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,809
; FILING DATE: December 18, 1995
; APPLICATION NUMBER: 60/023,943
; FILING DATE: August 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 223/113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-755-728-3

Query Match 2.0%; Score 7; DB 2;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLPPSAL 181
|||||
Db 334 VLPPSAL 340

RESULT 6

US-08-974-655-3
; Sequence 3, Application US/08974655
; Patent No. 5972676
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,655
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/755,728
; FILING DATE: No. 5972676ember 25, 1996
; APPLICATION NUMBER: 60/008,809
; FILING DATE: December 18, 1995
; APPLICATION NUMBER: 60/023,943
; FILING DATE: August 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 223/113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-974-655-3

Query Match 2.0%; Score 7; DB 2;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLPPSAL 181
|||||
Db 334 VLPPSAL 340

RESULT 7

US-09-283-011-3
; Sequence 3, Application US/09283011
; Patent No. 6207401
; GENERAL INFORMATION:
; APPLICANT: Florman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,011
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/012,135
; FILING DATE: January 22, 1998
; APPLICATION NUMBER: 08/755,728
; FILING DATE: No. 6207401ember 25, 1996
; APPLICATION NUMBER: 60/023,943
; FILING DATE: August 14, 1996
; APPLICATION NUMBER: 60/008,809
; FILING DATE: December 18, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 231/282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-283-011-3

Query Match 2.0%; Score 7; DB 4; Length 344;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLPPSAL 181
|||||
Db 334 VLPPSAL 340

RESULT 8

US-09-016-000-1
; Sequence 1, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,000
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0465 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 2940
US-09-016-000-1

Query Match 2.0%; Score 7; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLPPSAL 181
|||||
Db 337 VLPPSAL 343

RESULT 9
US-09-172-841-51
; Sequence 51, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621

; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 51
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-172-841-51

Query Match 2.0%; Score 7; DB 4; Length 456;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 TDVALEH 331
Db 357 TDVALEH 363

RESULT 10
US-08-672-571A-3
; Sequence 3, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-571A-3

Query Match 2.0%; Score 7; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 MPQFTEA 170
Db 43 MPQFTEA 49

RESULT 11
US-08-672-571A-1
; Sequence 1, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-571A-1

Query Match 2.0%; Score 7; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 MPQFTEA 170
Db 72 MPQFTEA 78

RESULT 12
US-08-559-492-12
; Sequence 12, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA

```
;
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR#154
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-12

Query Match 2.0%; Score 7; DB 2; Length 561;
Best Local Similarity 100.0%; Pred. No. 97; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 27 SSKFQFS 33
      |||||
Db 276 SSKFQFS 282

RESULT 13
US-08-426-819A-36
; Sequence 36, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318oml
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiko
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; TITLE OF INVENTION: Potentiator
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; US-08-426-819A-36

Query Match 2.0%; Score 7; DB 1; Length 584;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 180 ALPLDLL 186
      |||||
Db 79 ALPLDLL 85

RESULT 14
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 9174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..602
; OTHER INFORMATION: /note= "Represents residues 98 to
; OTHER INFORMATION: 731 of human granulocyte colony-stimulating
; OTHER INFORMATION: factor-receptor."
; US-08-419-652-6

Query Match 2.0%; Score 7; DB 2; Length 602;
```

Best Local Similarity 100.0%; Pred. No. le+02; 2.0%; Score 7; DB 1; Length 616;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Indels 0; Gaps 0;

QY 3 LFGLLLL 9
|||||||
Db 513 LFGLLLL 519

QY 36 KEONGVQ 42
|||||||
Db 35 KEONGVQ 41

RESULT 15

US-07-879-617A-13
; Sequence 13, Application US/07879617A
; Patent No. 5580775
; GENERAL INFORMATION:
; APPLICANT: Fremieu Jr., Robert T.
; APPLICANT: Caron, Marc G.
; APPLICANT: Blakely, Randy D.
; TITLE OF INVENTION: A High Affinity L-Proline Transporter
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/879,617A
; FILING DATE: 19920501
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU109

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rattus
; TISSUE TYPE: Brain - dopamine transporter

FEATURE:
; NAME/KEY: Active-site
; LOCATION: 99..100
; OTHER INFORMATION: /note= "Leucine zipper motif"

FEATURE:
; NAME/KEY: Active-site
; LOCATION: 106..107
; OTHER INFORMATION: /note= "Leucine zipper motif"

FEATURE:
; NAME/KEY: Active-site
; LOCATION: 113..114
; OTHER INFORMATION: /note= "Leucine zipper motif"

FEATURE:
; NAME/KEY: Active-site
; LOCATION: 120..121
; OTHER INFORMATION: /note= "Leucine zipper motif"

US-07-879-617A-13

Search completed: August 22, 2001, 14:31:11
Job time: 120 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2001, 14:29:11 ; Search time 20.99 Seconds
(without alignments)
996.440 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 345
Sequence: 1 MSFGLLLTSLAQKQGT.....DVALEHHECDVCRGRTGG 345

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	100.0	345	20	AA1980
2	345	100.0	345	20	AA1981
3	345	100.0	345	20	AA1982
4	345	100.0	345	21	AA1983
5	345	100.0	345	21	AA1984
6	345	100.0	345	21	AA1985
7	345	100.0	345	21	AA1986
8	345	100.0	345	21	AA1987
9	345	100.0	345	21	AA1988
10	345	100.0	345	21	AA1989
11	345	100.0	345	21	AA1990

12	345	100.0	345	21	AA19578
13	345	100.0	345	21	AA19579
14	345	100.0	345	21	AA19580
15	345	100.0	345	21	AA19581
16	345	100.0	345	21	AA19582
17	345	100.0	345	21	AA19583
18	345	100.0	345	21	AA19584
19	345	100.0	345	21	AA19585
20	345	100.0	345	21	AA19586
21	345	100.0	345	21	AA19587
22	327	91.9	318	21	AA19588
23	317	91.9	318	21	AA19589
24	312	90.4	354	21	AA19590
25	312	90.4	354	21	AA19591
26	306	88.7	339	21	AA19592
27	306	88.7	345	21	AA19593
28	306	88.7	345	21	AA19594
29	306	88.7	345	21	AA19595
30	245	71.0	282	21	AA19596
31	218	63.2	297	21	AA19597
32	154	44.6	167	21	AA19598
33	143	41.4	168	21	AA19599
34	116	33.6	149	21	AA19600
35	108	31.3	113	21	AA19601
36	108	31.3	113	21	AA19602
37	91	26.4	227	21	AA19603
38	91	26.4	227	21	AA19604
39	52	15.1	345	21	AA19605
40	52	15.1	345	21	AA19606
41	52	15.1	345	21	AA19607
42	24	7.0	180	21	AA19608
43	21	6.1	21	21	AA19609
44	21	6.1	21	21	AA19610
45	20	5.8	22	21	AA19611

ALIGNMENTS

RESULT 1

AA19578
ID AAY33679 standard; Protein; 345 AA.

XX AAY33679;

XX 11-JAN-2000 (first entry)

XX Human VEGF-E protein.

XX VEGF-E; human; vascular endothelial cell growth factor; wound repair;
treatment; cardiovascular disorder; endothelial disorder; therapy;
tissue generation; regeneration; cardiac hypertrophy; cancer; detection;
angiogenic disorder; age-related macular degeneration; vascular disease;
neovascularization; tumor; gene mapping.

XX Homo sapiens.

XX WO9947677-A2.

XX 23-SEP-1999.

XX 10-MAR-1999; 99WO-US05190.

XX 17-MAR-1998; 98US-0040220.

XX 02-NOV-1998; 98US-0184216.

XX (GETH) GENENTECH INC.

XX Ferrara N, Kuo SS;

XX WPI; 1999-580306/49.

XX N-PSDB; AAZ23691.

XX

Human PRO200 (vasc
Human PRO200 prote
Human PRO713 prote
Human TANGO 128.
Human growth facto
Bone morphogenic p
Human PRO200 prote
Human PRO200 prote
Human angiogenesis
Human VEGF-X prote
Amino acid sequenc
A fragment of plat
Human VEGF-X prote
Human VEGF-X prote
Lung cancer associ
Human VEGF-X homol
Human VEGF-X prote
Human VEGF-X prote
Human growth facto
Human VEGF-X prote
Bone morphogenic p
Human VEGF-X prote
Human VEGF-X CUB-1
Human VEGF-X PDGF-
Human VEGF-X prote
Human VEGF-X prote
Human VEGF-X prote
Mouse zveg3, SEQ
Murine vascular en
A murine platelet-
Murine TANGO 128.
Antigenic peptide
Antigenic peptide
Antigenic peptide

PT New growth factor polypeptide useful for treating cardiovascular or
 XX endothelial disorders, e.g. cardiac hypertrophy
 PS Claim 1; Fig 2; 122pp; English.
 XX
 CC This invention describes the isolation of a novel human vascular
 CC endothelial cell growth factor-E (VEGF-E) polypeptide which has
 CC tranquilizer, vulnery and cardiant activity. VEGF-E can be administered
 CC therapeutically, especially by expressing encoding polynucleotides, to
 CC treat cardiovascular or endothelial disorders in mammals, especially
 CC humans. It is useful in wound repair and tissue generation and
 CC regeneration, and may especially be used to treat cardiac hypertrophy
 CC It can be combined with a carrier in pharmaceutical compositions, which
 CC can be administered to treat disorders as above. VEGF-E can be used to
 CC screen for antagonists and agonists, and the antagonists administered to
 CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
 CC age-related macular degeneration. It can be used to generate antibodies,
 CC useful therapeutically as antagonists, as above. The antibodies are also
 CC useful to detect VEGF-E polypeptide, especially to diagnose
 CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
 CC by contacting the antibody with a tissue sample and detecting formation
 CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
 CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
 CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
 CC tissue samples. They can also be used to diagnose a disease or
 CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
 CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
 CC detecting a mutation in the VEGF-E-encoding sequence isolated from a
 CC sample. They may also be used to produce probes useful to detect related
 CC sequences or for gene mapping. This sequence represents the human VEGF-E
 CC protein described in the method of the invention.
 XX
 SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 20; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFLGILLTSLAQROGTQAESLSSKFQSSKNQGVODPOHERITVSTNGSIHS 60
 DB 1 msifglillttsalagrgrgtqaeslsskfssnkeqngvqdpqheritvtstngsihs 60
 QY 61 PRFPTYPRNTVLVRLVAVENVVIQTFDERFGLDEPDDICKYDFVEPEPSDGTIL 120
 DB 61 prfptyprntvlvrlvaveenvviqtfderfgleddedickdyfvevepsdgtll 120
 QY 121 GRWCGSGTVPGKQISKGNQIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPSVLPSPA 180
 DB 121 grwcsgstvpqkiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlpssa 180
 QY 181 LPDLNLNNAITAFSTLEDLIRLEPERWQLDEDLYRPTWQLLGRAVFGKRSRVVDLNL 240
 DB 181 lpdlnltnaitafstledlirleperwqldedlyrptwqlkgafvgrksrvvdlnl 240
 QY 241 LTEEVRVLSCTPRNSVSIRELKKRTDTIFWPGCLLVKRCGNCACCLHNCNECCVPSK 300
 DB 241 lteevrlyscptrnsvsireelkkrttdtifwpgcllvkrccgncacclhncnecvcpsk 300
 QY 301 VTKYHYEVLQRLPKTGVRLGHLKSLTDVALEHHECDVCVCRSTGG 345
 DB 301 vtkeyhevlqlrpkgtvrglghlksltdvalehhecdvcvcrstgg 345

RESULT 2

AAAY41766

ID AAY41766 standard; Protein; 345 AA.

XX AAY41766;

AC AAY41766;

XX 07-DEC-1999 (first entry)

DT 07-DEC-1999 (first entry)

XX

DE Human PRO200 protein sequence.

 XX Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KW secreted protein; transmembrane protein.
 XX Homo sapiens.
 OS WO9946281-A2.
 PN 16-SEP-1999.
 PD
 XX
 PF 08-MAR-1999; 99MO-US05028.
 XX
 PR 10-MAR-1998; 98US-0077450.
 PR 11-MAR-1998; 98US-0077632.
 PR 11-MAR-1998; 98US-0077641.
 PR 11-MAR-1998; 98US-0077649.
 PR 12-MAR-1998; 98US-0077791.
 PR 13-MAR-1998; 98US-0078004.
 PR 17-MAR-1998; 98US-0040220.
 PR 20-MAR-1998; 98US-0078886.
 PR 20-MAR-1998; 98US-0078910.
 PR 20-MAR-1998; 98US-0078936.
 PR 20-MAR-1998; 98US-0078939.
 PR 25-MAR-1998; 98US-0079294.
 PR 26-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079663.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 27-MAR-1998; 98US-0079728.
 PR 27-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 31-MAR-1998; 98US-0079923.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 31-MAR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 01-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081070.
 PR 08-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 09-APR-1998; 98US-0081229.
 PR 15-APR-1998; 98US-0081817.
 PR 15-APR-1998; 98US-0081838.
 PR 15-APR-1998; 98US-0081952.
 PR 15-APR-1998; 98US-0081955.
 PR 21-APR-1998; 98US-0082568.
 PR 21-APR-1998; 98US-0082569.
 PR 22-APR-1998; 98US-0082700.
 PR 22-APR-1998; 98US-0082704.
 PR 22-APR-1998; 98US-0082804.
 PR 23-APR-1998; 98US-0082767.
 PR 23-APR-1998; 98US-0082796.
 PR 27-APR-1998; 98US-0083336.
 PR 29-APR-1998; 98US-0083392.
 PR 29-APR-1998; 98US-0083495.
 PR 29-APR-1998; 98US-0083496.
 PR 29-APR-1998; 98US-0083499.
 PR 29-APR-1998; 98US-0083500.
 PR 29-APR-1998; 98US-0083545.
 PR 29-APR-1998; 98US-0083554.
 PR 29-APR-1998; 98US-0083558.
 PR 30-APR-1998; 98US-0083559.
 PR 05-MAY-1998; 98US-0083742.
 PR 06-MAY-1998; 98US-0084366.
 PR 06-MAY-1998; 98US-0084414.

PR 06-MAY-1998; 98US-0084441.
 PR 07-MAY-1998; 98US-0084598.
 PR 07-MAY-1998; 98US-0084600.
 PR 07-MAY-1998; 98US-0084627.
 PR 07-MAY-1998; 98US-0084637.
 PR 07-MAY-1998; 98US-0084639.
 PR 07-MAY-1998; 98US-0084640.
 PR 07-MAY-1998; 98US-0084643.
 PR 13-MAY-1998; 98US-0085323.
 PR 13-MAY-1998; 98US-0085338.
 PR 13-MAY-1998; 98US-0085339.
 PR 15-MAY-1998; 98US-0085573.
 PR 15-MAY-1998; 98US-0085579.
 PR 15-MAY-1998; 98US-0085580.
 PR 15-MAY-1998; 98US-0085582.
 PR 15-MAY-1998; 98US-0085689.
 PR 15-MAY-1998; 98US-0085697.
 PR 15-MAY-1998; 98US-0085700.
 PR 15-MAY-1998; 98US-0085704.
 PR 18-MAY-1998; 98US-0086023.
 PR 22-MAY-1998; 98US-0086392.
 PR 22-MAY-1998; 98US-0086414.
 PR 22-MAY-1998; 98US-0086430.
 PR 22-MAY-1998; 98US-0086486.
 PR 28-MAY-1998; 98US-0087098.
 PR 28-MAY-1998; 98US-0087106.
 PR 28-MAY-1998; 98US-0087208.
 PR 30-JUL-1998; 98US-0094651.
 PR 11-SEP-1998; 98US-0100038.
 PA (GETH) GENENTECH INC.
 XX
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX
 DR WPI; 1999-551358/46.
 DR N-PSDB; AAZ34296.
 XX
 PT New secreted and transmembrane polypeptides and their polynucleotides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders
 XX
 PS Claim 12; Fig 207; 530pp; English.
 XX
 CC The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to create transgenic
 CC animals. The proteins can be used to treat a variety of diseases and
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. AAZ33891 to
 CC AAZ34338, and AA41685 to AA41774 represent polynucleotide and
 CC polypeptide sequence given in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 345 AA;
 Query Match 100.0%; Score 345; DB 20; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLFGLLLTTSALAGQGTQAEENLSSKFFQSSNKQNGVQDPQHERIITVSTNGSIHS 60
 DB 1 mslfgllltsalagrqtqaeenlsskffqssnkngvqdpqheriitvstngsihs 60
 QY 61 PRFPPTPTNTVLVRLVAVENNVQLTDERFGLDEPDDICKYDFVEEPESDGTIL 120
 DB 61 prfphptptntvlvrlvaveenvqltderfgledpddickdyfveeepsdgtil 120
 QY 121 GRWCGSGTVPKQISKGNQIRIRFVSDEYFPSPGFCIHYNVMPQTEAVSPVLPPSA 180
 DB 121 grwcgsgtvpkgiskgnqirirfvsdeyfpsepgfcihynvmpqteavspvlppsa 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQLEDLYRPTWQLLKAFVFGKRSRVVDNL 240
 DB 181 lpldlnnaitafstledliryleperwqledlyrptwqlilgkafvfgkrsvvdl 240
 QY 241 LTEEVRVLSCTPRNFSVSIRRELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
 DB 241 lteevrlyscprnfsvsirreelkrttdtlfwpgcllvkrcgncaccclhncnecqcvpsk 300
 QY 301 VTKKYHEVLQRLPKTGVRLGKSLTDALEHHEEDCVCRCSTGG 345
 DB 301 vtkeyhevlqlrpkgtvrglghksltdvalehheecdcvrcgstgg 345
 RESULT 3
 AAY30023
 ID AAY30023 standard; protein; 345 AA.
 XX
 AC AAY30023;
 XX
 DT 11-OCT-1999 (first entry)
 XX
 DE Human vascular endothelial growth factor related protein.
 XX
 KW Vascular endothelial growth factor related protein; VEGF-R protein;
 KW tissue growth inhibition; tumour growth; cancer; tissue growth;
 KW angiogenesis; coronary artery blockage.
 XX
 OS Homo sapiens.
 XX
 PN WO9937671-A1.
 XX
 PD 29-JUL-1999.
 XX
 PF 26-JAN-1999; 99WO-US01574.
 XX
 PR 31-AUG-1998; 98US-0098548.
 PR 27-JAN-1998; 98US-0072635.
 PR 05-JUN-1998; 98US-0088089.
 PR 24-JUN-1998; 98US-0090544.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Dou S, Na S, Song HY;
 XX
 DR WPI; 1999-458680/38.
 DR N-PSDB; AAX86352.
 XX
 PT A vascular endothelial growth factor related protein and related
 PT polynucleotide, useful for identifying antagonists and binding
 PT compounds
 XX
 PS Claim 1; Page 56-58; 62pp; English.
 XX
 CC The present sequence represents a vascular endothelial growth factor
 CC related (VEGF-R) protein. VEGF-R can be used in assays to identify
 CC compounds that bind to it or that antagonize its activity. VEGF-R
 CC antagonists (e.g. anti-VEGF-R antibodies) are useful for inhibiting
 CC tissue growth. This is useful for inhibiting tumour growth and for
 CC treating cancer. VEGF-R itself can be used to stimulate tissue
 CC growth, angiogenesis and to treat coronary artery blockage. The
 CC VEGF-R coding sequence can be used for the recombinant production of
 CC the VEGF-R protein.
 XX
 SQ Sequence 345 AA;
 Query Match 100.0%; Score 345; DB 20; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLFGLLLTTSALAGQGTQAEENLSSKFFQSSNKQNGVQDPQHERIITVSTNGSIHS 60
 DB 1 mslfgllltsalagrqtqaeenlsskffqssnkngvqdpqheriitvstngsihs 60

Db 1 mslfllllltsalagrgqqtgaesnlsskfqssnkeqngvqdpqheriitvtstngsihs 60
 QY 61 PRPPTYPRNTVLVWRLVAEENVMWLTQTFDERFGLDEDDICKYDFVEVEPPSDGTIL 120
 Db 61 prfphtyprntvlwrlvaveenvwqltfderrfgledpeddickdydfvevepsdgtll 120
 QY 121 GRWCGSGVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNVMPQFTFVAVSPSVLPSSA 180
 Db 121 grwsgsgvpgkqiskngqirirfvsdeyfpsepgfcihynvmpqfteavspsvlpps 180
 QY 181 LPDLNLNNAITAFSTLEDLIRYLEPERWQDLEDLIRPTWQLLGKAFVGRKSRVVDNL 240
 Db 181 lpdlnltnaitafstledliryleperwqldledlirptwqllgkafvgrksrvvdnl 240
 QY 241 LTEEVRVLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
 Db 241 lteevrlyscptprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnccqvpsk 300
 QY 301 VTKKYHEVLQRLPKTGVGRGLHKSLLTDVALEHHEECDCVCRGSTGG 345
 Db 301 vtkeyhevlqlrpkgtgvgrglhkslldvalehheecdcvcrgstgg 345

RESULT 4

AAB48657
 ID AAB48657 standard; Protein; 345 AA.

XX AAB48657;

DT 09-MAR-2001 (first entry)

DE Human zveg3, SEQ ID NO:33.

XX Human; zveg3; zveg4 fusion; growth factor homologue; VEGF/PDGF family;
 KW CUB domain; PDGF-like activity; mitogenic; osteogenic;
 EW neovascularisation; tissue repair; proliferation; differentiation;
 KW liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
 KW periodontal disease; bone fracture; wound healing; vulnery; ischaemia;
 KW immunomodulation; hepatic.

XX Homo sapiens.

XX WO200066736-A1.

XX 09-NOV-2000.

XX 03-MAY-2000; 2000WO-US40047.

XX 03-MAY-1999; 99US-0304216.

XX 10-NOV-1999; 99US-0164463.

XX 04-FEB-2000; 2000US-0180169.

XX (ZYMO) ZYMOGENETICS INC.

XX Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;

XX WPI; 2000-687541/67.

XX N-PSDB; AAC81582.

XX Growth factor homologs and the nucleic acids that encode them, useful
 PT e.g. for treating liver damage, ischemia, multiple sclerosis and
 PT Alzheimer's disease -

XX Claim 48; Page 125-126; 143pp; English.

XX The invention relates to the human growth factor homologue zveg4
 CC (AAB48653), and nucleic acids encoding it (AAC81555). Zveg4 is a member
 CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
 CC growth factor) family. Zveg4 has a growth factor domain (AAB48654)
 CC characterised by a PDGF cysteine knot structure, and a CUB domain
 CC (AAB48655) which has a beta barrel structure. Zveg4 has PDGF-like
 CC activity, having mitogenic activity on fibroblasts, vascular smooth
 CC muscle cells and pericytes, and has also been shown to stimulate bone

CC growth. The invention also relates to fusion proteins comprising human
 CC zveg4 or fragments thereof, particularly human zveg4/human zveg3
 CC fusions; expression constructs and host cells comprising human zveg4
 CC nucleic acids; the recombinant expression of human zveg4; an antibody
 CC which binds to human zveg4 or a fragment thereof; a method of activating
 CC a cell-surface PDGF receptor using a zveg4-derived polypeptide; a
 CC method of modulating the proliferation, differentiation, migration or
 CC metabolism of bone cells, comprising exposing bone cells to
 CC zveg4-derived polypeptides; and a method of detecting a genetic
 CC abnormality in the zveg4 gene of a patient. Zveg4 proteins and derived
 CC fragments may be used to stimulate tissue development or repair, or
 CC cellular differentiation or proliferation. They are particularly used for
 CC the treatment or repair of liver damage, and may also be used to
 CC modulate neurite growth (e.g., in the treatment of Alzheimer's disease or
 CC multiple sclerosis). Due to their osteogenic activity, they may be used
 CC in the treatment of periodontal disease and fractures. They may also be
 CC used to enhance expansion and mobilisation of haematopoietic stem cells
 CC and endothelial precursor stem cells, which may be useful in the
 CC treatment of ischaemia, in wound healing, and in the modulation of the
 CC immune system. The present sequence represents human zveg3.
 XX
 SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLEGLLLLTALAGRQGTQAESNLSSKQFSSNKQNGVQDPQHERIITVSTNGSIHS 60
 Db 1 mslfglllltsalagrgqqtgaesnlsskfqssnkeqngvqdpqheriitvtstngsihs 60

QY 61 PRPPTYPRNTVLVWRLVAEENVMWLTQTFDERFGLDEDDICKYDFVEVEPPSDGTIL 120
 Db 61 prfphtyprntvlwrlvaveenvwqltfderrfgledpeddickdydfvevepsdgtll 120

QY 121 GRWCGSGVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNVMPQFTFVAVSPSVLPSSA 180
 Db 121 grwsgsgvpgkqiskngqirirfvsdeyfpsepgfcihynvmpqfteavspsvlpps 180

QY 181 LPDLNLNNAITAFSTLEDLIRYLEPERWQDLEDLIRPTWQLLGKAFVGRKSRVVDNL 240
 Db 181 lpdlnltnaitafstledliryleperwqldledlirptwqllgkafvgrksrvvdnl 240

QY 241 LTEEVRVLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
 Db 241 lteevrlyscptprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnccqvpsk 300

QY 301 VTKKYHEVLQRLPKTGVGRGLHKSLLTDVALEHHEECDCVCRGSTGG 345
 Db 301 vtkeyhevlqlrpkgtgvgrglhkslldvalehheecdcvcrgstgg 345

RESULT 5

AAB24250

ID AAB24250 standard; Protein; 345 AA.

XX AAB24250;

DT 08-FEB-2001 (first entry)

DE Human platelet-derived growth factor related protein LP8.

KW Human; platelet derived growth factor related protein; LP8; VEGFh;
 KW vascular endothelial growth factor h; tissue regeneration; vulnery;
 KW atherosclerosis; PDGF-related protein; antiarteriosclerotic.

XX Homo sapiens.

XX WO200059940-A2.

XX 12-OCT-2000.

XX

PF 24-MAR-2000; 2000WO-US06427.
XX
PR 06-APR-1999; 99US-0127913.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Hammond LJ, Na S;
XX
XX
DR WPI; 2000-664991/64.
DR N-PSDB; AAC64426.
XX
XX
PT Enhancing tissue growth and promoting wound healing by administering
PT platelet-derived growth factor related protein, LP8 or its analog and
PT treating atherosclerosis by administering LP8 antagonist
XX
XX
PS Claim 4; Page 63-64; 64pp; English.
XX
XX
CC The present invention describes a method for enhancing tissue growth,
CC promoting wound healing or stimulating smooth muscle growth by
CC administering a platelet-derived growth factor (PDGF) related protein,
CC designated LP8 or its analogue. Also described is a method of slowing
CC the progress of atherosclerosis or treating atherosclerosis comprising
CC the administration of an LP8 antagonist. The method is useful for
CC enhancing tissue growth, promoting wound healing and stimulating smooth
CC muscle growth. Antagonists of LP8 are useful for treating
CC atherosclerosis. The present sequence represents human LP8, which is
CC also called VEGFh.
XX
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFLGILLTTSALAGROGTQAESNLSSKFQFSSNKQNGVODPOHERIITVSTNGSIHS 60
Db 1 mslfgillttsalagrqgtqaesnlsskfqfssnkqngvqdpqheriitvstngsihs 60
QY 61 PRFPHYPRNTVLVRLVAVENWVWIOITDFERFGLDEPDICKYDFVEEPEPSDGTIL 120
Db 61 prfphyprntvlvrlvaveenwvwlitdfderfgledpeddickdyfveeepsdgtil 120
QY 121 GRWCGSGTVPGKOISKGNOIRIRFVSDYFPEPFCGCIHYNVMPQTEAVSPVLPPSA 180
Db 121 grwcsgstvgpkqiskgnqirirfvsdeyfpsepfgciahynvmpqfteavspvlppsa 180
QY 181 LPDLILNNAITAFSTLEDLIRYLEPERWQDLEDLIRPTWQLLGKAFVGRKSRVDNL 240
Db 181 lpdlilnnaitafstledliryleperwqldledlirptwqllgkatfgrksrvdnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREEELKRTDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPSK 300
Db 241 lteevrlyscprnfsvsiireelkrtdtifwpgcllvkrccgncacclhncnccqvpsk 300
QY 301 VTKKYHEVLQRLPKTGVGRGLHSLTDVALEHHEECDCVCRGSGTG 345
Db 301 vtkkyhevlqlrpkgtvgrihksltdvalehheecdcvcrsgtg 345

RESULT 6
ID AAB44322
XX AAB44322 standard; Protein; 345 AA.
XX AC AAB44322;
XX
DT 08-FEB-2001 (first entry)
XX Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
DE Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
XX Human expressed sequence tag; detection; cancer.
KW
XX

OS Homo sapiens.
XX WO200053756-A2.
XX
XX 14-SEP-2000.
XX
XX 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 21-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin LJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX
XX WPI; 2000-611443/58.
DR N-PSDB; AAC78582.
XX
XX Novel PRO polypeptides and polynucleotides used in detection methods,
XX to target bioactive molecules to specific cells, and to modulate
XX cellular activities -
XX
XX Claim 12; Fig 207; 636pp; English.
XX
XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed
XX sequence tag) sequences which encode secreted or transmembrane PRO
XX polypeptides. The PRO polynucleotides and polypeptides have cytostatic
XX activity. The polynucleotides and polypeptides can be used for detecting
XX the presence of PRO polypeptides in samples, for linking bioactive
XX molecules to cells and for modulating biological activities of cells,
XX using the polypeptides for specific targeting. The polypeptide targeting
XX can be used to kill the target cells, e.g. for the treatment of cancers.
XX The polypeptide pairs provide specific targeting of bioactive molecules
XX to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
XX the isolation of the PRO polynucleotide sequences.
XX
XX Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFLGILLTTSALAGROGTQAESNLSSKFQFSSNKQNGVODPOHERIITVSTNGSIHS 60
Db 1 mslfgillttsalagrqgtqaesnlsskfqfssnkqngvqdpqheriitvstngsihs 60
QY 61 PRFPHYPRNTVLVRLVAVENWVWIOITDFERFGLDEPDICKYDFVEEPEPSDGTIL 120
Db 61 prfphyprntvlvrlvaveenwvwlitdfderfgledpeddickdyfveeepsdgtil 120
QY 121 GRWCGSGTVPGKOISKGNOIRIRFVSDYFPEPFCGCIHYNVMPQTEAVSPVLPPSA 180
Db 121 grwcsgstvgpkqiskgnqirirfvsdeyfpsepfgciahynvmpqfteavspvlppsa 180

Db 121 grwsgtvpqkqskngqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180

QY 181 LPDLNNATAFSTLEDLIRYLEPERWQDLEDLRPTWQLLGKAFVFGKSRVVDLNL 240

Db 181 lpdllnnatfstledliryleperwqldledlirptwqllgkafvfgkrsvvdlnl 240

QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGLLVKRCGGNCACCLHNCNCCQVPSK 300

Db 241 lteevrlyscptprnfsvsireelkrttdtifwpgccllvkrccgncacclhncnccqcvpsk 300

QY 301 VTKKYHEVLQRLPKTGVRCGLHLSLTDVALEHHEECDCVCRGSTGG 345

Db 301 vtckyhevlqlrpkgtgvrghlsldvalehheecdcvcrgstgg 345

RESULT 7

AAB10633

ID AAB10633 standard; Protein; 345 AA.

XX AC AAB10633;

XX DT 19-JAN-2001 (first entry)

XX DE Human RACE generated VEGF-X protein.

XX KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;

XX KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;

XX KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;

XX KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

XX KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;

XX KW venous sore; diabetic ulcer; burns; skin graft growth.

XX OS Homo sapiens.

XX PN WO200037641-A2.

XX PD 29-JUN-2000.

XX PF 21-DEC-1999; 99WO-US30503.

XX PR 22-DEC-1998; 98GB-0028377.

XX PR 18-MAR-1999; 99US-0124967.

XX PR 08-NOV-1999; 99US-0164131.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;

XX PI Dhanaraj SN, Xu J;

XX DR WPI: 2000-442669/38.

XX DR N-PSDB; AAA71951.

XX PT New vascular endothelial growth factor protein, useful for treating or

XX PT preventing diseases associated with inappropriate angiogenesis activity

XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -

XX PS Disclosure; Fig 6; 127pp; English.

XX CC This invention describes a novel vascular endothelial growth factor-X

XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has

XX CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and

XX CC antidiabetic activity and acts as an angiogenesis and vascularization

XX CC regulator. An antisense molecule of the invention is useful for treating

XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic

XX CC retinopathy by inhibiting angiogenic activity or inappropriate

XX CC vascularization including formation and proliferation of new blood

XX CC vessels, growth and development of tissues, tissue regeneration and organ

XX CC and tissue repair in a subject. The products of the invention are useful

XX CC for preparing medicaments for treating wounds such as dermal ulcers,

XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote

XX CC skin graft growth, tissue repair, proliferation of new blood vessels,

XX CC tissue regeneration and organ repair by promoting angiogenic activity or

XX CC vascularization. This sequence represents the RACE generated human VEGF-X

CC protein described in the method of the invention.

XX SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSIFGILLTTSALAGORQGTQAESNLSSKFQFSNKKFQNGVQDPOHERIITVSTNGSIHS 60

Db 1 msifgillttsalagrqgtqaesnlsskfqsnnkeqngvqdpqherilitvstngsihs 60

QY 61 PRFPHYPTNTVLVWRLVAVEENVWLIQLTDERFGLDEPDDICKYDFVEVEPDSGTIL 120

Db 61 prfphyptntvlvwrivaveenvwliqtderfgledpeddickdyfvevepsdgtill 120

QY 121 GRWCGSTVPKGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPVLPSSA 180

Db 121 grwcgstvpqkqskngqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180

QY 181 LPDLNNATAFSTLEDLIRYLEPERWQDLEDLRPTWQLLGKAFVFGKSRVVDLNL 240

Db 181 lpdllnnatfstledliryleperwqldledlirptwqllgkafvfgkrsvvdlnl 240

QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGLLVKRCGGNCACCLHNCNCCQVPSK 300

Db 241 lteevrlyscptprnfsvsireelkrttdtifwpgccllvkrccgncacclhncnccqcvpsk 300

QY 301 VTKKYHEVLQRLPKTGVRCGLHLSLTDVALEHHEECDCVCRGSTGG 345

Db 301 vtckyhevlqlrpkgtgvrghlsldvalehheecdcvcrgstgg 345

RESULT 8

AAB10635

ID AAB10635 standard; Protein; 345 AA.

XX AC AAB10635;

XX DT 19-JAN-2001 (first entry)

XX DE Human VEGF-X protein #1 isolated from clones 4 and 7.

XX KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;

XX KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;

XX KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;

XX KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

XX KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;

XX KW venous sore; diabetic ulcer; burns; skin graft growth.

XX OS Homo sapiens.

XX PN WO200037641-A2.

XX PD 29-JUN-2000.

XX PF 21-DEC-1999; 99WO-US30503.

XX PR 22-DEC-1998; 98GB-0028377.

XX PR 18-MAR-1999; 99US-0124967.

XX PR 08-NOV-1999; 99US-0164131.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;

XX PI Dhanaraj SN, Xu J;

XX DR WPI: 2000-442669/38.

XX DR N-PSDB; AAA71955.

XX PT New vascular endothelial growth factor protein, useful for treating or

XX PT preventing diseases associated with inappropriate angiogenesis activity

PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX Disclosure; Fig 9; 127pp; English.
 XX This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGF-X protein
 CC isolated from clones 4 and 7 described in the method of the invention.
 XX Sequence 345 AA;
 SQ

Query Match 100.0%; Score 345; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTSLALAGROGTOAESNLSSKFQSSNKEQNGVODPOHERIIVSTNGSIHS 60
 DB 1 msflgillttsalagrggtqaesnlsskfqsnskeqngvqdpqheriitvstngsihs 60
 QY 61 PRFPTYPRNTVLVWRLVAEENVMWQLTDFRFGLEDDEDDICKYDFVEEPEPSDGTIL 120
 DB 61 prfptyprntvlvwrilvaeenvmwqltderfgleddedickdyfveeepsdgtl 120
 QY 121 GRWCGSGTVPGKQISGNQIRIRFVSDEYFPSEPGFCIHYNVMPQFTEAVSPVLPPSA 180
 DB 121 grwcgsgtvpkgiskgnqirirfvsdeyfpsepgfcihynvmpqfteavspvlppsa 180
 QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVGRKSRVDNLN 240
 DB 181 lpdlnnaitafstledliryleperwqldledlyrptwqlkafvgrksrvdnl 240
 QY 241 LTEEVRYSCTPRNFSVSIREEELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPSPK 300
 DB 241 lteevrlyscprnfsvsiireelkrttdtlfwpgcllvkrccgncacclhncnecqcvp 300
 QY 301 VTKKYHEVLQRLPKTGVGRGLHKSITDVALEHHEECDCVCRGSTGG 345
 DB 301 vtkkyhevlqlrpkgtvgrglhksltdvalehheecdcvcrgstgg 345

RESULT 9
 AAB10644
 ID AAB10644 standard; Protein; 345 AA.
 XX
 AC AAB10644;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human VEGF-X protein #4.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; tissue repair; proliferation of new blood vessels;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.
 XX
 OS Homo sapiens.
 XX
 PN WO200037641-A2.

XX 29-JUN-2000.
 XX 21-DEC-1999; 99WO-US30503.
 XX 22-DEC-1998; 98GB-0028377.
 PR 18-MAR-1999; 99US-0124967.
 PR 08-NOV-1999; 99US-0164131.
 XX (JANC) JANSSEN PHARM NV.
 XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 XX WPI; 2000-442669/38.
 DR N-PSDB; AAA71990.
 XX New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX Disclosure; Fig 30B; 127pp; English.
 XX This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents a human VEGF-X protein
 CC described in the method of the invention.
 XX Sequence 345 AA;
 SQ

Query Match 100.0%; Score 345; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTSLALAGROGTOAESNLSSKFQSSNKEQNGVODPOHERIIVSTNGSIHS 60
 DB 1 msflgillttsalagrggtqaesnlsskfqsnskeqngvqdpqheriitvstngsihs 60
 QY 61 PRFPTYPRNTVLVWRLVAEENVMWQLTDFRFGLEDDEDDICKYDFVEEPEPSDGTIL 120
 DB 61 prfptyprntvlvwrilvaeenvmwqltderfgleddedickdyfveeepsdgtl 120
 QY 121 GRWCGSGTVPGKQISGNQIRIRFVSDEYFPSEPGFCIHYNVMPQFTEAVSPVLPPSA 180
 DB 121 grwcgsgtvpkgiskgnqirirfvsdeyfpsepgfcihynvmpqfteavspvlppsa 180
 QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVGRKSRVDNLN 240
 DB 181 lpdlnnaitafstledliryleperwqldledlyrptwqlkafvgrksrvdnl 240
 QY 241 LTEEVRYSCTPRNFSVSIREEELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPSPK 300
 DB 241 lteevrlyscprnfsvsiireelkrttdtlfwpgcllvkrccgncacclhncnecqcvp 300
 QY 301 VTKKYHEVLQRLPKTGVGRGLHKSITDVALEHHEECDCVCRGSTGG 345
 DB 301 vtkkyhevlqlrpkgtvgrglhksltdvalehheecdcvcrgstgg 345

RESULT 10

AAB10650
 ID AAB10650 standard; Protein; 345 AA.
 AC AAB10650;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human 990126veg protein.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.
 XX
 OS Homo sapiens.
 XX
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 PR 18-MAR-1999; 99US-0124967.
 PR 08-NOV-1999; 99US-0164131.
 XX
 XX (JANC) JANSSEN PHARM NV.
 XX
 PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 XX
 DR WPI; 2000-442669/38.
 XX
 PT New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX
 PS Disclosure; Fig 11; 127pp; English.
 XX
 CC This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human 990126veg protein
 CC used to illustrate the method of the invention.
 XX
 XX Sequence 345 AA;
 SQ
 Query Match 100.0%; Score 345; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFLGALLLSALAGROGTOAESNLSSKQFSSNKEONGVDPQHERIITVSTNGSIHS 60
 DB 1 mslflllltsalagrgtgaesnlsskqfssnkeqngvdpqheriitvstngsihs 60
 QY 61 PRFHTYPRNTVLVRLVAVENWVQITFDRFRGLEDPEDDICKYDFVEEPPSDGTL 120
 DB 61 prfhtyprntvlvrlvaveenwvqitfdrfrgledpeddickdyfveeppsdgtl 120

QY 121 GRWCGSTVPGKQISKQNIQIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPVLPPSA 180
 DB 121 grwcgstvpqkiskqniqirfvsdeyfpsepgfcihynivmpqteavspvlppsa 180
 QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDNL 240
 DB 181 lpdlldlnnaitafstledliryleperwqdlledlyrptwqllgkatvgrksrvvdnl 240
 QY 241 LTEVRLYSCYTPNFSVSIREELKRTDTIFWPGCLLVKRGCGNACCLHMCNECQCVPSPK 300
 DB 241 lteevrlyscytpnfsvsireelkrtdtifwpgcllvkrgcgncacclhmcneqcvpstk 300
 QY 301 VTKKYHEVLRLPKPTVGRGLHKSJTDVALEHHEPCDCVCRGSTGG 345
 DB 301 vtckyhevllrpkptvgrglhksjtdvalehhepcdcvcrgstgg 345
 RESULT 11
 AAB10651
 ID AAB10651 standard; Protein; 345 AA.
 AC AAB10651;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human VEGF-X protein #3.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.
 XX
 OS Homo sapiens.
 XX
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 PR 18-MAR-1999; 99US-0124967.
 PR 08-NOV-1999; 99US-0164131.
 XX
 XX (JANC) JANSSEN PHARM NV.
 XX
 PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 XX
 DR WPI; 2000-442669/38.
 XX
 PT New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX
 PS Claim 72; Fig 12; 127pp; English.
 XX
 CC This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human 990126veg protein
 CC used to illustrate the method of the invention.

CC vascularization. This sequence represents the human VEGF-x protein
described in the method of the invention.

XX Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGLLLTSALAGORQGTQAESNLSSKFKQSNKQNGVQDPQHERIITVSTNGSIHS 60
|||||
Db 1 msifgllltsalagrggtqaesnlsskfkqsnkqngvqdpqheriitvstngsihs 60
|||||

QY 61 PRFPHYPRNTVLVWRLVAVEENWVWQLTDFERFGLDEPDICKYDFVEVEEPSDGTIL 120
|||||
Db 61 prfphyprntvlwrlvaveenwvwltdferfgledpeddickdyfveveepsdgtil 120
|||||

QY 121 GRWCGSGTVPGKQISKGNQIRIRFVSDYFSPBPGFCIHYNVMPQTEAVSPVLPPSA 180
|||||
Db 121 grwcsgstvpqkqiskgnqirirfvsdypfbpgfcihynvmppqteavspvlppsa 180
|||||

QY 181 LPDLNNALATSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
|||||
Db 181 lpdlnnalatastledliryleperwqdlledlyrptwllgkafvgrksrvvdlnl 240
|||||

QY 241 LTEVRLYSCTPRNFSVIREELKRTDTIFWPGCLLVKRCGNCACCLHNCQCVPFSK 300
|||||
Db 241 lteevrlyscprnfsvireelkrttdtlfwpgcllvkrcgncacclhncnecqcvpsk 300
|||||

QY 301 VTKKYHEVLQRPKTGVRGLHKSITDVALEHHEBCDCVCRGSTGG 345
|||||
Db 301 vtckyhevqlrpkgtgvrghksitdvalhehheecdcvcrgstgg 345
|||||

RESULT 12

AAB19578
ID AAB19578 standard; Protein; 345 AA.

XX AC AAB19578;

XX 22-JAN-2001 (first entry)

XX Human PRO200 (vascular endothelial growth factor E).

XX PRO200; vascular epithelial growth factor E; VEGF-E; human;
KW ocular disease; retinopathy; maculopathy; therapy;
KW retinitis pigmentosa; macular degeneration; retinal detachment;
KW retinal tear; macular hole; myopia; traumatic choriorretinopathy;
KW acute retinal necrosis syndrome; contusion; edema;
KW retinal vision occlusion; vascular disease; retinal vasculitis;
KW thrombocytopenic purpura; uveitis; retinal occlusion.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..14

FT Protein /label= Signal_peptide

FT Protein 15..345

FT Modified-site /label= Mature_Pro200

FT Modified-site 25..29

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 55..59

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 254..258

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 15..21

FT Modified-site /note= "N-myristoylation"

FT Modified-site 117..123

FT Modified-site /note= "N-myristoylation"

FT Modified-site 127..133

FT Modified-site /note= "N-myristoylation"

FT Modified-site 281..287

FT Modified-site /note= "N-myristoylation"
FT 282..288
FT Modified-site /note= "N-myristoylation"
FT 319..325
FT Modified-site /note= "Amidation"

XX WO200053760-A2.

XX 14-SEP-2000.

XX 10-MAR-2000; 2000WO-US06319.

XX 12-MAR-1999; 99US-0123957.

XX (GETH) GENENTECH INC.

XX Ferrara N, Goddard A, Gurney AL, Hebert C, Henzel WJ, Kabakoff RC;

XX Klein RD, Kljavin LJ, Kuo SS, La Fleur M, Wood WI;

XX WPI; 2000-587437/55.

XX DR N-ESDB; AAA88515.

XX Novel PRO polypeptides useful for preventing or rescuing retinal cells
from injury caused by ocular diseases such as retinitis pigmentosa,
retinopathy, retinal degenerative diseases, degenerative myopia,
uveitis

XX Claim 2; Fig 2; 140pp; English.

XX The present sequence is that of human PRO200 or vascular
endothelial growth factor E (VEGF-E), as predicted from a cDNA
clone (see AAA88515) that was isolated from a glioma cell line G61
library using probes (see AAA8523-26) based on an expressed sequence
tag (see AAA88522) that showed homology to VEGF. PRO200 has a
predicted mol. wt. of 39,029 and a pI of about 6.06. A method for
producing PRO polypeptides, including PRO200, using a host cell
transformed with a vector comprising a PRO nucleic acid is claimed.
The invention relates to the use of PRO polypeptides to delay,
prevent or rescue retinal cells such as retinal neurons selected from
photoreceptors, retinal ganglion cells, displaced retinal ganglion
cells, amacrine cells, displaced amacrine cells, horizontal and
bipolar neurons, and supportive cells (including Mueller cells and
pigment epithelial cells) from injury and degradation. The retinal
cells are preferably photoreceptors and photoreceptor cell injury or
death is caused by retinal injury, light or environmental trauma or
by an ocular disease selected from retinitis pigmentosa, macular
degeneration, including age-related, retinal detachment, retinal
tears, retinopathy, retinal degenerative diseases, macular holes,
degenerative myopia, acute retinal necrosis syndrome, traumatic
choriorretinopathies or contusion such as Purtscher's retinopathy,
edema, ischemic conditions such as central or branch retinal vision
occlusion, collagen vascular diseases, thrombocytopenic purpura,
uveitis, retinal vasculitis and occlusion associated with Eales
disease and systemic lupus erythematosus (claimed).

XX Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGLLLTSALAGORQGTQAESNLSSKFKQSNKQNGVQDPQHERIITVSTNGSIHS 60

Db 1 msifgllltsalagrggtqaesnlsskfkqsnkqngvqdpqheriitvstngsihs 60

QY 61 PRFPHYPRNTVLVWRLVAVEENWVWQLTDFERFGLDEPDICKYDFVEVEEPSDGTIL 120

Db 61 prfphyprntvlwrlvaveenwvwltdferfgledpeddickdyfveveepsdgtil 120

QY 121 GRWCGSGTVPGKQISKGNQIRIRFVSDYFSPBPGFCIHYNVMPQTEAVSPVLPPSA 180

Db 121 grwcsgstvpqkqiskgnqirirfvsdypfbpgfcihynvmppqteavspvlppsa 180

QY 181 LPDLNNATAFSTLEDLIRYLEPERWOLDLEDLYRPTWQLLKGAFVFGKSRVVLDNL 240
 |||||||
 Db 181 lpldlinnaftstledliryleperwqldledlyrptwqlilgkafvfgkrsrvvldnl 240
 |||||||
 QY 241 LTEEVRVLYSCTPRNFSVIREELKRTDTLFWPGCLLVKRCGNCACCLHNCNQCVPK 300
 |||||||
 Db 241 lteevrlyscprnfsvsireelkrttdtlfwpgccllvkrccgncacclhncnecqcvpsk 300
 |||||||
 QY 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEEDCVCRGSGTG 345
 |||||||
 Db 301 vtkkyhevlqlrpkgtvrglghksltdvalehheecdcvcrgstgg 345
 |||||||

RESULT 13
 AAB33414
 ID AAB33414 standard; Protein; 345 AA.
 XX
 AC AAB33414;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Human PRO200 protein UNQ174 SEQ ID NO:2.
 XX
 KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;
 KW dermatological; antiarthritic; antirheumatic; immunosuppressive;
 KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
 KW antianaemic; hepatotropic; virucide; antiposrotic; antiallergic;
 KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
 KW autoimmune thrombocytopaenia; immune-mediated renal disease;
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;
 KW inflammatory bowel disease; gluten-sensitive enteropathy;
 KW autoimmune disease; immune-mediated skin disease; allergic disease;
 KW immunological disease; transplantation associated disease;
 KW graft rejection; graft-versus-host-disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200053758-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05841.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99WO-US08615.
 PR 28-APR-1999; 99US-0131445.
 PR 04-MAY-1999; 99US-0132371.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.

PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
 XX
 DR WPI: 2000-572271/53.
 DR N-PSDB; AAC58579.
 XX
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 XX
 PS Claim 33; Fig 2; 309pp; English.
 XX
 CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 XX
 SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLEGLLLTSALAGQGTOAESNLSSKFOFSNKEQNGVDPQHERITVTNSGHS 60
 |||||||
 Db 1 msllglllltsalagrqtgaesnlsskfssnkeqngvqdpqheritvtnsghs 60
 |||||||
 QY 61 PRPPTYPRNTVLVRLVAVENNVWVQLTDFERFGLDEPDDICKYDFVEVEPSDGTIL 120
 |||||||
 Db 61 prfphtyprntvlvrlvaveenvwqltfdferfgldedpddickdyfvevepsdgtll 120
 |||||||
 QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFPSPGPGCIHYNVNPOFTAVSPSVLPSPA 180
 |||||||
 Db 121 grwcgsgtvpkgqiskngnqirirfvsdeyfpsepggfcihynvmpdftavspsvlpssa 180
 |||||||
 QY 181 LPDLNNATAFSTLEDLIRYLEPERWOLDLEDLYRPTWQLLKGAFVFGKSRVVLDNL 240
 |||||||
 Db 181 lpldlinnaftstledliryleperwqldledlyrptwqlilgkafvfgkrsrvvldnl 240
 |||||||
 QY 241 LTEEVRVLYSCTPRNFSVIREELKRTDTLFWPGCLLVKRCGNCACCLHNCNQCVPK 300
 |||||||

```

Db 241 lteevlyscprnfsireelkrtidtfwpgcllvkrcggncacclhncnecqcvpsk 300
QY 301 VTKKYHEVLQRPKTGVRGLHKSITDVALEHHECDVCVCRGSTGG 345
Db 301 vtkkyhevlqrpktgvrghksitdvalehhecdvcvcrgstgg 345

RESULT 14
AAB24412
ID AAB24412 standard; Protein; 345 AA.
AC AAB24412;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human PRO713 protein sequence SEQ ID NO:137.
XX
KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;
KW cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200032221-A2.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US28313.
XX
PR 01-DEC-1998; 98WO-US25108.
PR 16-DEC-1998; 98US-0112850.
PR 12-JAN-1999; 99US-0115554.
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
XX
PA (GETH ) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Ferrara M, Gerber H, Hillan KJ, Goddard A;
PI Godowski PU, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Watanabe CK, Williams PM, Wood WI;
XX
DR WPI; 2000-412154/35.
DR N-PSDB; AAA77621.
XX
PT Nucleic acids encoding PRO polypeptides useful for preventing,
PT diagnosing and treating a cardiovascular, endothelial or
PT angiogenic disorders in mammals .
XX
PS Claim 72; Fig 50; 315pp; English.
XX
CC The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating a cardiovascular
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,
CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or

```

```

CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGQROGTQAESNLSSKFOFSSNKQNGVODPOHERIIVSTNGSIHS 60
Db 1 mslfgllltsalagrggtqaesnlsskfgfssnkeqngvdpqcherliitvstngsihs 60

QY 61 PRFHTYPRNTVLVRLVAVENWVQLTDFEREGLEDDEDDICKYDFVEVEEPSDGTIL 120
Db 61 prfhtyprntvlvrlvaveenwvqltderfegledpeddickdyfveveepsdgtill 120

QY 121 GRWCGSGTVPKGKQISKNQIRIRFVSDYFDPSEPGFCIHYNIVMPQFTAVSPVLPESA 180
Db 121 grwcsgtvpkgkqisknqirirfvsdyyfdpsepgfcihynivmpqfteavspvlpesa 180

QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLIRFTWLLGKAFVFGKRSRVVDLNL 240
Db 181 lpdllnnaitafstledliryleperwqldledlirftwllgkafvfgkrsvvdlnl 240

QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDIFWPGLLVRCGGNCACCLHNCNCCQCVPSK 300
Db 241 lteevrlyscprnfsvsireelkrtidifwpgcllvrcggncacclhncnecqcvpsk 300

QY 301 VTKKYHEVLQRPKTGVRGLHKSITDVALEHHECDVCVCRGSTGG 345
Db 301 vtkkyhevlqrpktgvrghksitdvalehhecdvcvcrgstgg 345

RESULT 15
AAB01419
ID AAB01419 standard; Protein; 345 AA.
AC AAB01419;
XX
DT 20-OCT-2000 (first entry)
XX
DE Human TANGO 128.
XX
KW TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human.
XX
OS Homo sapiens.
XX
PN WO200039284-A1.
XX
PD 06-JUL-2000.
XX
PF 23-DEC-1999; 99WO-US31025.
XX
PR 30-DEC-1998; 98US-0223546.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA;
XX
DR WPI; 2000-465743/40.

```

DR N-PSDB; AAA47452.
XX
PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases
XX
PS Claim 8; Fig 1; 209pp; English.
XX
CC Nucleic acids encoding TANGO polypeptides are useful as modulating
CC agents for regulating cellular processes like asthma, graft
CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
CC lupus erythematosus. The nucleic acids are also useful for producing
CC transgenic animals and the TANGO polypeptides themselves. Partial
CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
CC forensic biology, for diagnostic assays, prognostic assays,
CC pharmacogenomics and for monitoring clinical trials. TANGO
CC polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a
CC disorder associated with aberrant TANGO expression. A wide range
CC of cellular disorders can be treated.
XX
SQ Sequence 345 AA:

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLLGILLTSLAGORQGTQAESNLSSKFFQSSNKEQNGVODPOHERITVTSTNGSIHS 60
Db 1 msllgillttsalagrggtqaesnlsskfqsnskeqngvqdpqheritvtstngsihs 60

Qy 61 PRFPTYPRNTVLVRLVAEENWVLIQTDFDERFGLDEPDDEDDICKYDFVEVEEPSDGTIL 120
Db 61 prfptyprntvlvrlvaveenvwliqtderfgleddedddickdyfveveepsdgtil 120

Qy 121 GRWCGSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPOFTAVSPSVLPSPA 180
Db 121 grwcsgstvpgkqiskgnqirirfvsdeyfpsepgfcihynivmpoftavspsvlppsa 180

Qy 181 LPDLNLNNAITAFSTLEDLIRYLEPERWQOLDLEDLYRPTWQLLGKAFVGRKSRVVDNL 240
Db 181 lpdlnltnaitafstledliryleperwqoldledlyrptwqlllgkafvgrksrvvndl 240

Qy 241 LTEEVRLYSCTPRNFSVSIRELKRKRTDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPSK 300
Db 241 lteevrlyscptrnfsvsirelkrkrtdtifwpgcllvkrccgncacclhncnccqvpsk 300

Qy 301 VTKKYHEVLQLRPKTGVRGLHKSITDVALEHHEEDCVCVRGSGTG 345
Db 301 vtkeyhevlqlrpkgtvrglhlksitdvalhheecdvcvrgstgg 345

Search completed: August 22, 2001, 14:31:38
Job time: 147 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2001, 14:31:02 ; Search time 12.76 seconds
(without alignments)
926.187 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 345

Sequence: 1 MSFLGLLLLSALAGRQGT.....DVALEHHEECDCVCRGSTGG 345

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	8	2.3	707	1 BMP1_XENLA	P98070 xenopus lae
2	8	2.3	986	1 BMP1_HUMAN	P13497 homo sapien
3	8	2.3	991	1 BMP1_MOUSE	P98063 mus musculus
4	7	2.0	75	1 LEUO_SALTY	P46924 salmonella
5	7	2.0	148	1 UMP1_YEAST	P38293 saccharomyc
6	7	2.0	261	1 Y612_SYNY3	P72581 synechocyst
7	7	2.0	313	1 YHCL_BACSU	P54593 bacillus su
8	7	2.0	314	1 LEUO_ECOLI	P10151 escherichia
9	7	2.0	340	1 CGGR_BACSU	O32533 bacillus su
10	7	2.0	343	1 NTDO_MOUSE	O61327 mus musculus
11	7	2.0	377	1 DCA2_DIACA	P33213 rhizobium f
12	7	2.0	384	1 AMPC_PROST	O39677 dianthus ca
13	7	2.0	393	1 KRL_VZVD	O69773 providencia
14	7	2.0	419	1 CPET_SYNPY	P09251 varicella-z
15	7	2.0	423	1 YCDB_ECOLI	Q02174 synechococc
16	7	2.0	471	1 NOLX_RHIFR	P31545 escherichia
17	7	2.0	477	1 PEN3_ADECC	P33213 rhizobium f
18	7	2.0	489	1 PEN3_ADEML	O69550 canine aden
19	7	2.0	513	1 ARO9_YEAST	O10439 mouse adeno
20	7	2.0	513	1 NHAB_ECOLI	P38840 saccharomyc
21	7	2.0	525	1 NCAP_RINDL	P27377 escherichia
22	7	2.0	557	1 CO9_RABIT	P37708 rinderpest
23	7	2.0	578	1 OAR2_LYMTS	P48747 oryctolagus
24	7	2.0	596	1 NOLX_RHISN	O01670 lymaea sta
25	7	2.0	608	1 HSCA_BUCAP	P55711 rhizobium s
26	7	2.0	619	1 NTDO_RAT	O51883 buchnera ap
27	7	2.0	620	1 NTDO_HUMAN	P23977 rattus norv
28	7	2.0	628	1 MSLN_HUMAN	O01959 homo sapien
29	7	2.0	693	1 NTDO_BOVIN	O13421 homo sapien
30	7	2.0	835	1 INVA_YEREN	P27922 bos taurus
31	7	2.0	836	1 GCSR_HUMAN	P19196 yersinia en
32	7	2.0	1039	1 YAF3_SCHPO	O99062 homo sapien
33	7	2.0	1049	1 ITA5_HUMAN	O09857 schizosacch
					P08648 homo sapien

34	7	2.0	1053	1 ITA5_MOUSE	P11688 mus musculus
35	7	2.0	3321	1 KEND_HUMAN	O95613 homo sapien
36	6	1.7	88	1 C552_MARRY	P82903 marinobacte
37	6	1.7	91	1 GONI_PIG	P49921 sus scrofa
38	6	1.7	95	1 NQOB_THETH	Q56226 thermus aqu
39	6	1.7	97	1 YHBY_ECOLI	P42550 escherichia
40	6	1.7	98	1 YPSB_BACSU	P50839 bacillus su
41	6	1.7	102	1 CYC_EUGVI	P22342 euglena vir
42	6	1.7	102	1 MGN_ORYSA	P49030 oryza sativ
43	6	1.7	103	1 YA40_MYCPN	P75074 mycoplasma
44	6	1.7	105	1 RL31_AERPE	Q9y25 aeropyrum p
45	6	1.7	106	1 GUAV_MOUSE	O09051 mus musculus

ALIGNMENTS

RESULT 1

ID	BMP1_XENLA	STANDARD;	PRT;	707 AA.
AC	P98070;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.-) (BMP-1).			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo;			
RX	MEDLINE=94085787; PubMed=8262384;			
RA	Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;			
RT	"Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic protein-1 during early embryonic development.";			
RL	Gene 134:257-261(1993).			
CC	FUNCTION: INVOLVED IN PATTERN FORMATION IN GASTRULA AND LATER DIFFERENTIATION OF DEVELOPING ORGANS.			
CC	DEVELOPMENTAL STAGE: BLASTULA, EARLY GASTRULA AND HATCHED TADPOLES; LITTLE OR NO EXPRESSION IN MORULA AND LATE GASTRULA.			
CC	SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	SIMILARITY: CONTAINS 3 CUB DOMAINS.			
CC	SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; L12249; AAA16313.1; --			
DR	HSSP; P00736; IAPQ.			
DR	MEROPS; M12.005; --			
DR	InterPro; IPR000130; --			
DR	InterPro; IPR000152; --			
DR	InterPro; IPR000561; --			
DR	InterPro; IPR000859; --			
DR	InterPro; IPR001506; --			
DR	InterPro; IPR001881; --			
DR	Pfam; PF01400; Astacin; 1.			
DR	Pfam; PF00431; CUB; 3.			
DR	Pfam; PF00008; EGF; 1.			
DR	PRINTS; PR00480; ASTACIN.			
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.			
DR	PROSITE; PS01180; CUB; 3.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.			
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.			
DR	PROSITE; PS01186; EGF_2; 1.			

DR PROSITE; PS01187; EGF_CA; 1.
 KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase; Protease;
 KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP 83 ? POTENTIAL.
 FT CHAIN 84 707 BONE MORPHOGENETIC PROTEIN 1.
 FT DOMAIN 84 284 METALLOPROTEASE.
 FT DOMAIN 285 397 CUB.
 FT DOMAIN 398 509 CUB.
 FT DOMAIN 510 551 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 554 666 CUB.
 FT METAL 176 176 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 177 177 BY SIMILARITY.
 FT METAL 180 180 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 186 186 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 514 526 BY SIMILARITY.
 FT DISULFID 522 535 BY SIMILARITY.
 FT DISULFID 537 550 BY SIMILARITY.
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC9B8D CRC64;

Query Match 2.38; Score 8; DB 1; Length 707;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111
 | | | | | | | |
 DB 607 CKYDFVEV 614

RESULT 2
 BMDP1_HUMAN STANDARD; PRT; 986 AA.
 ID BMDP1_HUMAN PRT; 986 AA.
 AC P13497; Q13292; Q99421; Q99422; Q99423; Q14874;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BONE MORPHOGENETIC PROTEIN 1; PRECURSOR (EC 3.4.24.19) (BMP-1)
 DE (PROCOLLAGEN C-PROTEINASE) (PCP) (MAMMALIAN TOLLOID PROTEIN) (MTLD).
 GN BMDP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM BMP1-3).
 RC TISSUE=Skin;
 RX PubMed=8643539;
 RA Li S.W., Sieron A.L., Fertala A., Hojima Y., Arnold W.V.,
 RA Prockop D.J.;
 RT "The C-proteinase that processes procollagens to fibrillar collagens
 RT is identical to the protein previously identified as bone morphogenic
 RT protein-1";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5127-5130(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM BMP1-1).
 RX MEDLINE=89072730; PubMed=3201241.
 RA Wozney J.M., Rosen V., Celeste A.J., Mitsuoka L.M., Whitters M.J.,
 RA Kriz R.W., Hewick R.M., Wang E.A.;
 RT "Novel regulators of bone formation: molecular clones and
 RT activities";
 RL Science 242:1528-1534(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS BMP1-4; BMP1-5 AND BMP1-6).
 RC TISSUE=Placenta;
 RX PubMed=9500680;
 RA Janitz M., Heiser V., Boettcher U., Landt O., Lauster R.;

RT "Three alternatively spliced variants of the gene coding for the human
 RT bone morphogenetic protein-1";
 RT J. Mol. Med. 76:141-146(1998).
 RN [4]
 RP SEQUENCE OF 703-986 FROM N.A. (ISOFORM BMP1-3).
 RC TISSUE=Placenta;
 RX MEDLINE=95096114; PubMed=7798260;
 RA Takahara K., Lyons G.E., Greenspan D.S.;
 RT "Bone morphogenetic protein-1 and a mammalian tolloid homologue (mtld)
 RT are encoded by alternatively spliced transcripts which are
 RT differentially expressed in some tissues";
 RT J. Biol. Chem. 269:32572-32578(1994).
 RL -!- FUNCTION: CLEAVES THE C-TERMINAL PROPEPTIDES OF PROCOLLAGEN I, II
 CC AND II. INDUCES CARTILAGE AND BONE FORMATION.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF THE C-TERMINAL PROPEPTIDE AT
 CC ALA-|-ASP IN TYPE I AND II PROCOLLAGENS AND AT ARG-|-ASP IN TYPE
 CC III.
 CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
 CC -!- ENZYME REGULATION: ACTIVITY IS INCREASED BY THE PROCOLLAGEN C-
 CC ENDOPEPTIDASE ENHANCER PROTEIN.
 CC -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; BMP1-1, BMP1-2, BMP1-3 (SHOWN
 CC HERE), BMP1-4, BMP1-5 AND BMP1-6; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -!- TISSUE SPECIFICITY: UBQUITOUS.
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 5 CUB DOMAINS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
 CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U50330; AAA93462.1; -
 CC EMBL; M22488; AAA51833.1; -
 CC EMBL; Y08723; CAA69973.1; -
 CC EMBL; Y08724; CAA69974.1; -
 CC EMBL; Y08725; CAA69975.1; -
 CC EMBL; L35279; AAC41710.1; -
 CC PIR; A37278; A37278.
 CC HSSP; P00736; 1APQ.
 CC MEROPS; M12.005; -
 CC MIM; I12264; -
 CC InterPro; IPR000130; -
 CC InterPro; IPR000152; -
 CC InterPro; IPR000561; -
 CC InterPro; IPR000859; -
 CC InterPro; IPR001506; -
 CC InterPro; IPR001881; -
 CC Pfam; PF01400; Astacin; 1.
 CC Pfam; PF00431; CUB; 5.
 CC Pfam; PF00008; EGF; 2.
 CC PRINTS; PR00480; ASTACIN
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS01180; CUB; 5.
 CC PROSITE; PS00010; ASX_HYDROXYL; 2.
 CC PROSITE; PS00022; EGF_1; FALSE_NEG.
 CC PROSITE; PS01186; EGF_2; 2.
 CC PROSITE; PS01187; EGF_CA; 2.
 KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase;
 KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
 KW Glycoprotein; Zymogen; Alternative splicing.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 120 POTENTIAL.
 FT CHAIN 121 986 BONE MORPHOGENETIC PROTEIN 1.
 FT DOMAIN 121 321 METALLOPROTEASE.
 FT DOMAIN 322 434 CUB 1.
 FT DOMAIN 435 546 CUB 2.
 FT DOMAIN 547 588 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).


```
FT DISULFID 579 592 BY SIMILARITY.
FT DISULFID 596 622 BY SIMILARITY.
FT DISULFID 649 671 BY SIMILARITY.
FT DISULFID 712 723 BY SIMILARITY.
FT DISULFID 719 732 BY SIMILARITY.
FT DISULFID 734 747 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 991 AA; 111607 MW; 68A1847783A0BB9E CRC64;

Query Match 2.3%; Score 8; DB 1; Length 991;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 KYDFVEV 111
    |||||
Db 649 KYDFVEV 656

RESULT 4
LEUO_SALTY
ID LEUO_SALTY STANDARD; PRT; 75 AA.
AC P45924;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE ACTIVATOR PROTEIN IN LEUABCD OPERON (FRAGMENT).
GN LEUO.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86223773; PubMed=3519576;
RA Haughn G.W., Wessler S.R., Gemunill R.M., Calvo J.M.;
RT "High A + T content conserved in DNA sequences upstream of leuABCD in
RL Escherichia coli and Salmonella typhimurium.";
RN J. Bacteriol. 166:1113-1117(1986).
[2]
CC CONCEPTUAL TRANSLATION.
RA Bairoch A.;
RL Unpublished observations (OCT-1995).
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
CC INTRODUCED IN POSITION 5 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12892; AAB02429.1; ALT_FRAME.
CC StyGene; SG10657; leuO.
CC InterPro; IPR000847; -.
CC Pfam; PF00126; HTH_1; 1.
CC PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Transcription regulation; DNA-binding.
FT DNA_BIND 39 58 H-T-H MOTIF (BY SIMILARITY).
FT NON_TER 75 75
SQ SEQUENCE 75 AA; 8507 MW; 42F2E678796762CD CRC64;

* Query Match 2.0%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 2.0%; Score 7; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 STNGSIH 59
    |||||
Db 100 STNGSIH 106

RESULT 6
Y612_SYNY3
ID Y612_SYNY3 STANDARD; PRT; 261 AA.
```

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 VDLNLLT 242
    |||||
Db 22 VDLNLLT 28

RESULT 5
UMPL_YEAST
ID UMP1_YEAST STANDARD; PRT; 148 AA.
AC P38293;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEASOME MATURATION FACTOR UMP1.
GN UMP1 OR YBR173C OR YBR1234.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YHP500;
RX MEDLINE=98150854; PubMed=9491890;
RA Ramos P.C., Hoeckendorff J., Johnson E.S., Varshavsky A., Dohmen R.J.;
RT "Ump1 is required for proper maturation of the 20S proteasome and
RT becomes its substrate upon completion of the assembly.";
RL Cell 92:489-499(1998).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
RA Niedemann E., Schenk-Groeninger R., Servos J., Wehner E.,
RA Wolter R., Brendel M., Bauer J., Braun H., Dern K., Duesterhus S.,
RA Gruenbein R., Hedges D., Kiesau P., Korol S., Krebs B., Proft M.,
RA Siegers K., Baur A., Boles E., Miosga T.,
RA Schaeff-Gerstenschlaeger I., Zimmermann F.K.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SHORT-LIVED CHAPERONE PRESENT IN THE PRECURSOR FORM OF
CC THE 20S' PROTEASOME AND ABSENT IN THE MATURE COMPLEX. REQUIRED FOR
CC THE CORRECT ASSEMBLY AND ENZYMIC ACTIVATION OF THE PRE2 PROPEPTIDE.
CC ALSO PREVENTS PREMATURE PROCESSING OF THE PRE2 PROPEPTIDE.
CC -!- MISCELLANEOUS: SEEMS TO BE DEGRADED BY THE PROTEASOME UPON ITS
CC FORMATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ002557; CAA05556.1; -.
CC EMBL; Z36042; CAA85134.1; -.
CC PIR; S46044; S46044.
CC SGD; S0000377; UMP1.
KW Proteasome; Chaperone.
SQ SEQUENCE 148 AA; 16760 MW; 7C774DD40F3FCD7C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 STNGSIH 59
    |||||
Db 100 STNGSIH 106

RESULT 6
Y612_SYNY3
ID Y612_SYNY3 STANDARD; PRT; 261 AA.
```

AC P72581;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE HYPOTHETICAL 21.0 KDA PROTEIN SLR0612.
 GN SLR0612.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=11448;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome, and assignment of potential protein-coding regions.";
 DNA Res. 3:109-136(1996).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Bairoch A.;
 RL Unpublished observations (SEP-1998).
 CC -!- SIMILARITY: BELONGS TO FAMILY I OF PSEUDOURIDINE SYNTHASES.
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
 CC INTRODUCED IN POSITION 20 TO PRODUCE THIS ORF.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D90899; BAA16580.1; ALT_FRAME.
 DR InterPro; IPR000613; -;
 DR InterPro; IPR000748; -;
 DR Pfam; PF00849; Pseudou_synth_2; 1.
 DR PROSITE; PS01149; PSI_RSU; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 261 AA; 29561 MW; 601A453085C04A69 CRC64;
 Query Match 2.0%; Score 7; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 GULLLTS 11
 DB 86 GULLLTS 92
 RESULT 7
 YHCD YHCD YHCD STANDARD; PRT; 313 AA.
 AC P54593;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HYPOTHETICAL 34.9 KDA PROTEIN IN GLPD-CSPB INTERGENIC REGION.
 GN YHCD
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124185; PubMed=8969498;
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;

RT "A 22 kb DNA sequence in the cspB-glppPKD region at 75 degrees on the
 RT Bacillus subtilis chromosome";
 RL Microbiology 142:3021-3026(1996).
 RN [2]
 RP SEQUENCE OF 303-313 FROM N.A.
 RC STRAIN=168 / JH642;
 RA Wendrich T.M., Marahiel M.A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X96983; CAA65692.1; -;
 DR EMBL; U58859; -; NOT_ANNOTATED_CDS.
 DR EMBL; Z99108; CAB12737.1; -;
 DR Subtilist; BG11587; yhcI.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 16 36 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 SQ SEQUENCE 313 AA; 34881 MW; 4020879A507F176D CRC64;
 Query Match 2.0%; Score 7; DB 1; Length 313;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LFGLLLL 9
 DB 154 LFGLLLL 160
 RESULT 8
 LEUO_ECOLI
 ID LEUO_ECOLI STANDARD; PRT; 314 AA.
 AC P10151; P75640;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROBABLE ACTIVATOR PROTEIN IN LEUABCD OPERON.
 GN LEUO.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88320486; PubMed=3413113;
 RA Henikoff S., Haughn G.W., Calvo J.M., Wallace J.C.;
 RT "A large family of bacterial activator proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6602-6606(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE OF 1-74 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=86223773; PubMed=3519576;
 RA Haugbo G.W., Wessler S.R., Gemmill R.M., Calvo J.M.;
 RT "High A + T content conserved in DNA sequences upstream of leuABCD in
 RT *Escherichia coli* and *Salmonella typhimurium*.";
 RL J. Bacteriol. 166:1113-1117(1986).
 CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M21150; AAA85299.1; -;
 DR EMBL; X55034; CAA38853.1; ALT INIT.
 DR EMBL; D10483; BAA01344.1; ALT INIT.
 DR EMBL; AE000118; AAC73187.1; ALT INIT.
 DR EMBL; M12891; AAA83880.1; ALT INIT.
 DR PIR; A29846; QQEC33.
 DR PIR; S14419; S14419.
 DR PIR; S14389; S14389.
 DR ECO2DBASE; F035.0; 6TH EDITION.
 DR EcoGene; EG10531; leuo.
 DR InterPro; IPR000847; -;
 DR Pfam; PF00126; HTH_1; 1.
 DR Pfam; PF01046; NodD_C term; 1.
 DR PRINTS; PR00039; HTHLYSR.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 KW Transcription regulation; DNA-binding.
 FT DNA_BIND 39 58 H-T-H MOTIF (BY SIMILARITY).
 FT CONFLICT 291 314 MISSING (IN REF. 1 AND 2).
 SQ SEQUENCE 314 AA; 35695 MW; 25605FA08C1073E3 CRC64;

 QY 236 VDLNLT 242
 Db 22 VDLNLT 28

 RESULT 9
 CGGR_BACSU STANDARD; PRT; 340 AA.
 ID CGGR_BACSU
 AC O32253;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CENTRAL GLYCOLYTIC GENES REGULATOR.
 GN CGGR.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 OC *Bacillus*/Staphylococcus group; *Bacillus*.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: REPRESSES THE TRANSCRIPTION OF THE GAPA GENE.
 CC -!- SIMILARITY: BELONGS TO THE SORC FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z99121; CAB15400.1; -;
 DR Subtilist; BG14085; cggR.
 KW Transcription regulation; DNA-binding; Repressor.
 FT DNA_BIND 37 56 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 340 AA; 37382 MW; 18C885966DB42DB CRC64;

 QY 68 PRNTVLV 74
 Db 321 PRNTVLV 327

 RESULT 10
 NTDO_MOUSE STANDARD; PRT; 343 AA.
 ID NTDO_MOUSE
 AC Q61327; Q60719;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE SODIUM-DEPENDENT DOPAMINE TRANSPORTER (DA TRANSPORTER) (DAT)
 DE (FRAGMENT).
 GN SLC6A3 OR DAT1 OR DAT.
 OS *Mus musculus* (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=95364623; PubMed=7637582;
 RA Donovan D.M., Vandenbergh D.J., Perry M.P., Bird G.S., Ingersoll R.,
 RA Nanthakumar E., Uhl G.R.;
 RT "Human and mouse dopamine transporter genes: conservation of
 RT 5'-flanking sequence elements and gene structures.";
 RL Brain Res. Mol. Brain Res. 30:327-335(1995).
 CC -!- FUNCTION: TERMINATES THE ACTION OF DOPAMINE BY ITS HIGH
 CC AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
 CC FAMILY (SNF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U15791; AAA86462.1; -;
 DR EMBL; U12313; AAA86462.1; JOINED.
 DR EMBL; U16265; AAC52283.1; -;
 DR MGD; MGI:94862; Dat1.
 DR InterPro; IPR000175; -;

FT	CHAIN	1	83	S-ADENOSYLMETHIONINE DECARBOXYLASE BETA CHAIN (BY SIMILARITY).
FT	CHAIN	84	377	S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA CHAIN (BY SIMILARITY).
FT	SITE	83	84	CLEAVAGE (NONHYDROLYTICAL) (BY SIMILARITY).
FT	MOD_RES	84	84	CONVERTED TO A PIRUVOYL GROUP (BY SIMILARITY).
FT	ACT_SITE	24	24	IMPORTANT FOR CATALYTIC ACTIVITY (BY SIMILARITY).
FT	ACT_SITE	27	27	IMPORTANT FOR CATALYTIC ACTIVITY (BY SIMILARITY).
FT	ACT_SITE	98	98	IMPORTANT FOR CATALYTIC ACTIVITY (BY SIMILARITY).
FT	SEQUENCE	377 AA;	41344 MW;	C3E58AAD143F1AD6 CRC64;
Query Match 2.08; Score 7; DB 1; Length 377;				
Best Local Similarity 100.0%; Pred. No. 32;				
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	235	VVDNLNL 241		
DB	280	VVDNLNL 286		
RESULT 12				
ID	AMPC_PROST	STANDARD;	PRT;	384 AA.
AC	069773;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	BETA-LACTAMASE PRECURSOR (EC 3.5.2.6) (CEPHALOSPORINASE).			
GN	AMPC.			
OS	Providencia stuartii.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Providencia.			
OX	NCBI_TaxID=588;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=VDC 96;			
RA	Koeck J.L., Basmaciogullari S., Parzy D., Barnaud G., Teyssou R.,			
RA	Buisson Y., Philippon A., Arlet G.J.;			
RT	"Cloning and sequencing of ampc and ampr genes from Providencia stuartii.";			
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: THIS PROTEIN IS A SERINE BETA-LACTAMASE WITH A SUBSTRATE SPECIFICITY FOR CEPHALOSPORINS.			
CC	-1- CATALYTIC ACTIVITY: A BETA-LACTAM + H(2)O = A SUBSTITUTED BETA-AMINO ACID.			
CC	-1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE CLASS-C BETA-LACTAMASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
DR	EMBL: Y17315; CAA76739.1; -			
DR	InterPro: IPR001466; -			
DR	InterPro: IPR001586; -			
DR	Pfam: PF001144; beta-lactamase; 1.			
DR	PROSITE; PS00336; BETA_LACTAMASE.C; FALSE_NEG.			
KW	Hydrolase; Antibiotic resistance; Periplasmic; Signal.			
FT	SIGNAL	1	29	POTENTIAL.
FT	CHAIN	30	384	BETA-LACTAMASE.
FT	ACT_SITE	87	87	BY SIMILARITY.
FT	BT_BINDING	337	339	SUBSTRATE (BY SIMILARITY).
SQ	SEQUENCE	384 AA;	443366 MW;	8CCACAF9B1377E9 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 384;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 DLIRYLE 204
 |||||
 Db 252 DLIRYLE 258

RESULT 13

KRL_VZVD
 ID KRL_VZVD STANDARD; PRT; 393 AA.
 AC P09251;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE SERINE/THREONINE-PROTEIN KINASE (EC 2.7.1.-).
 GN 66.
 OS Varicella-zoster virus (strain Dumas) (VZV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86306657; PubMed=3018124;
 RA Davison A.J., Scott J.E.;
 RT "The complete DNA sequence of varicella-zoster virus.";
 RL J. Gen. Virol. 67:1759-1816(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84131932; PubMed=6321154;
 RA Davison A.J.;
 RT "DNA sequence of the US component of the varicella-zoster virus genome.";
 RL EMBO J. 2:2203-2209(1983).
 RN [3]
 RP REVIEW.
 RX MEDLINE=89369695; PubMed=2855680;
 RA Leader D.P., Purves F.C.;
 RT "The herpesvirus protein kinase: a new departure in protein phosphorylation.";
 RL Trends Biochem. Sci. 13:244-246(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; X04370; CAA27949.1; -
 EMBL; X00208; CAA25031.1; -
 PIR; E27345; TVBE66.
 HSP; Q16539; 1WFC.
 InterPro: IPR000719; -
 InterPro: IPR002290; -
 Pfam: PF00069; pkinase; 1.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 Transferase; Serine/threonine-protein kinase; ATP-binding.
 DOMAIN 93 378 PROTEIN KINASE.
 NP_BIND 99 107 ATP (BY SIMILARITY).
 BINDING 122 122 ATP (BY SIMILARITY).
 ACT_SITE 206 206 BY SIMILARITY.
 SQ SEQUENCE 393 AA; 43679 MW; 2396280DC40AFBF7 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 393;
 Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 AGORQGT 20
 |||||
 Db 123 AGORQGT 129

RESULT 14

CPEY_SYPY
 ID CPEY_SYPY STANDARD; PRT; 419 AA.
 AC O02174;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE BILIN BIOSYNTHESIS PROTEIN CPEY.
 GN CPEY.
 OS Synechococcus sp. (strain WH8020).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=32052;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93144698; PubMed=8425055;
 RA de Lorimer R., Wilbanks S.M., Glazer A.N.;
 RT "Genes of the R-phycoerythrin II locus of marine Synechococcus spp.,
 RT and comparison of protein-chromophore interactions in phycocyanins
 RT differing in bilin composition.";
 RL Plant Mol. Biol. 21:225-237(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93123238; PubMed=8419325;
 RA Wilbanks S.M., Glazer A.N.;
 RT "Rod structure of a phycoerythrin II-containing phycobilisome. I.
 RT Organization and sequence of the gene cluster encoding the major
 RT phycobiliprotein rod components in the genome of marine Synechococcus
 RT sp. WH8020.";
 RL J. Biol. Chem. 268:1226-1235(1993).
 CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF BILIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M95288; AAA27337.1; -
 DR PIR; D45045; D45045.
 DR PIR; S31052; S31052.
 KW Phycobilisome.
 SQ SEQUENCE 419 AA; 47401 MW; 77A724FF8B42C55E CRC64;

Query Match 2.0%; Score 7; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 VVDLNL 241
 |||||
 Db 220 VVDLNL 226

RESULT 15

YCDB_ECOLI
 ID YCDB_ECOLI STANDARD; PRT; 423 AA.
 AC P31545; P75903;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 46.8 KDA PROTEIN IN PUTP-PHOX INTERGENIC REGION PRECURSOR
 DE (ORF1).
 GN YCDB.
 OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map";
RL DNA Res. 3:137-155(1996).
RN [3]
RP SEQUENCE OF 196-423 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=93186698; PubMed=8444794;
RA Kim S.-K., Makino K., Amemura M., Shinagawa H., Nakata A.;
RT "Molecular analysis of the phoH gene, belonging to the phosphate
regulon in Escherichia coli";
RL J. Bacteriol. 175:1316-1324(1993).
CC -!- SIMILARITY: TO B.SUBTILIS YWBN (IPA-29D).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000203; AAC74104.1; -;
DR EMBL; D90738; BAA35796.1; -;
DR EMBL; D90739; BAA35800.1; -;
DR EMBL; D10391; BAA01229.1; -;
DR PIR; A47065; A47065.
DR EcoGene; EG11735; YcDb.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.
KW Hypothetical protein; Signal.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 423 HYPOTHETICAL PROTEIN YCDB.
SQ SEQUENCE 423 AA; 46754 MW; 65D381F829DB2570 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 FDERFGL 96
|||||||
DB 143 FDERFGL 149

Search completed: August 22, 2001, 14:32:51
Job time: 109 sec